DEPARTMENT OF HEALTH & HUMAN SERVICES **Public Health Service**

Food and Drug Administration

Center for Biologics Evaluation and Research
1401 Rockville Pike
Rockville, MD 20852

Memorandum

Date: February 9, 1999

From: James Kaiser, M.D.

Through: Marc Walton, M.D.

Through: Karen Weiss, M.D.

Subject: Medical Officer's clinical review
BLA 98-0261
Serono Laboratories, Inc.
Interferon-β1a (Rebif®)

To: The File

Attached document:

Memo regarding MRI evaluations in clinical trial --------- by Dr. Robert Grossman of the University of Pennsylvania, dated October 19, 1998

TABLE OF CONTENTS

ABBREVIATIONS

multiple sclerosis	MS
central nervous system	CNS
relapsing-remitting multiple sclerosis	RRMS
secondary progressive multiple sclerosis	SPMS
magnetic resonance imaging	MRI
Expanded Disability Scale Score	EDSS
interferon	IFN
three times a week	TIW
subcutaneous	SC
intramuscular	IM
gadolinium	Gd

SYNOPSIS OF APPLICATION

This application, submitted on February 27, 1998, is for approval of Rebif, an interferon-β1a manufactured by Ares Serono, for the treatment of patients with relapsing-remitting ---------- ---------- MS. The proposed package insert states the following:

> "Rebif® (interferon-beta-1a) is indicated for the treatment of patients with relapsing-remitting multiple sclerosis ----------. Rebif® decreases the frequency ---------- of clinical exacerbations and delays ---------- accumulation of physical disability. Safety and efficacy in patients with progressive multiple sclerosis have not been established.."

Two dose regimens are being proposed: IFN-β1a at 22 mcg subcutaneously three times a week and 44 mcg subcutaneously three times a week.

BACKGROUND

MULTIPLE SCLEROSIS

Clinical considerations/epidemiology

MS is thought to be due to autoimmune destruction of central nervous system (CNS) myelin. It commonly starts with localized symptoms related to disorders of the optic nerves, spinal cord, and cerebral white matter. Localized pathology often becomes widely scattered throughout the white matter of the CNS, with disparate symptoms of visual, motor, sensory, and psychiatric nature. The natural course of the disease is highly variable. The most common clinical pattern is for the disease to start with a relapsing/remitting course, with accumulation of lesions that correlates with an evolution to secondary progressive disease. The frequency of relapses tends to decrease over time. In many patients a steady neurologic and physical deterioration occurs over a period of 30 to 40 years due to an accumulation of fixed neurological deficits. During the period between these two stages the disease has an intermediate character, and is referred to by some investigators as relapsing/progressive MS. Primary progressive MS, in which progression of deficit accumulation occurs from the onset without relapses, is seen in a minority of patients.

MS is the most common of the demyelinating disorders, with a prevalence of approximately 1 per 1000 persons in the United States and Europe. MS affects about 1 million young adults worldwide. There are over 250,000 patients with multiple sclerosis in the United States, with an annual incidence of approximately 9,000. Approximately 2-3 times as many women as men are affected. Approximately 2/3 of cases begin between the ages of 20-40 years.

The therapy of multiple sclerosis has changed in recent years with the approval of interferon-βs (see below), which are widely considered immune modulators, and copolymer 1. Corticosteroids are widely used for the treatment of exacerbations. Other more investigational immune modulators are azathioprine and cyclophosphamide. Symptomatic therapies include amantadine for treatment of fatigue, baclofen and benzodiazepines for spasticity, urologic antispasmodics for bladder disorders, and benzodiazepines and antidepressants for psychological disorders.

Diagnosis/measurement of disability

The diagnosis of MS generally requires confirmation of occurrence of two lesions that must have occurred at different times in different anatomical regions of the CNS. The lesions are usually demonstrable on magnetic resonance imaging (MRI). A "T1-weighted" MRI performed after the infusion of gadolinium (Gd) is believed to show cranial lesions of acute onset, the contrast agent leaking through the normally impermeable endothelial barrier. These lesions may resolve over a period of months. "T2-weighted" lesions are believed to represent fixed, residual pathology.

The predominant tool to measure the accumulation of disability is the expanded disability scale score (EDSS), which is determined by assessing the Kurtzke Functional Systems in each of 6 neurological areas (pyramidal, cerebellar, brainstem, sensory, bowel and bladder, and visual). EDSS scores range from 0 (normal) to 10 (death) in 1/2-unit steps. Patients are fully ambulatory through EDSS 4.5, after which progressive impairment in ambulation becomes the predominating factor in the EDSS.

INTERFERONS

The interferons are generally classified as types I and II, the former containing the α, β, and τ interferons, the latter consisting of γ interferon. Each type is coded by a distinct gene. Of the previously approved interferons for MS, one (Avonex, made by Biogen) is a β-1a interferon, and the other (Betaseron, made by Berlex) is a β-1b interferon. While a cytokine and cellular protein response profile to the administration of interferons has been characterized (for example, increased TNF-α and 2,5 OAS synthetase generation, antiviral and cellular antiproliferative activity), the mechanism of therapeutic effect of interferon-βs in MS is not known. Further, it is unclear what relationship the specific activity of interferons, which is expressed as antiviral activity against an NIH standard, has to its activity in treating multiple sclerosis.

In clinical use, the most common reported side-effects of administration of interferons are the flu-like symptoms fever, chills, headache, fatigue, asthenia, myalgia, and anorexia. Hematological (lymphopenia, neutropenia, thrombocytopenia, and anemia) and hepatic (AST/ALT) toxicities are known side-effects of interferon therapy.

Betaseron was the first interferon approved for the treatment of MS. The pivotal trial for Betaseron recruited 372 subjects with relapsing-remitting MS, with EDSS from 0-5.5. The product reduced the frequency of exacerbations by about 30%. The severity of exacerbations was reduced and the times to 1st and 2nd exacerbations were prolonged; it did not have an effect on disability. Betaseron's label states that it "is indicated for use in ambulatory patients relapsing-remitting multiple sclerosis to reduce the frequency of clinical exacerbations."

The pivotal trial for Avonex (Biogen), the second interferon approved for the treatment of MS, recruited 301 subjects with relapsing forms of MS and EDSS from 1-3.5. Avonex significantly delayed progression in disability and reduced the number of exacerbations at 2 years but did not influence the time to the first exacerbation nor the number of exacerbation-free subjects. It "is indicated for the treatment of relapsing forms of multiple sclerosis to slow the accumulation of physical disability and decrease the frequency of clinical exacerbations."

PRODUCT INFORMATION

Rebif is a syringe pre-filled with a liquid formulation of IFN-β1a. The interferon in Rebif is a glycoprotein produced in recombinant mammalian (Chinese Hamster Ovary or CHO) cells, with an amino acid sequence identical to that of fibroblast-derived human interferon-β. Rebif is formulated as a solution with human albumin, mannitol, sodium acetate, and water for injection.

SUMMARY OF CLINICAL TRIALS OF REBIF IN MS

Serono has completed 2 clinical trials in MS: ----------, a 72-subject, open-label trial of subjects with RRMS, and ----------, a 560-subject, placebo-controlled trial of subjects with RRMS that is the focus of this review. The primary endpoint of ---------- was MRI evidence of brain lesions; the primary endpoint of ---------- was exacerbations of MS. Trial ---------- was completed after ---------- was initiated.

DOSE SELECTION FOR PLACEBO-CONTROLLED TRIAL

Dose selection for the placebo-controlled trial ---------- was based upon Serono's assessment of tolerability of the 22 mcg and 44 mcg doses of their IFN-β1a in non-MS indications, as well as their assessment of Rebif's expected ratio of benefit to adverse effects in comparisons with Avonex and Betaseron.

SCOPE OF REVIEW

The focus of this review will be on ----------, as this constitutes the bulk of evidence for the efficacy and safety of Rebif in RRMS. Results from the 72-subject open-label trial ---------- will be summarized, but both its open-label design and the small number of subjects investigated make its contribution to the understanding of Rebif weak. Therefore, an integrated summary of efficacy will not be presented. Results from the integrated summary of safety, including adverse event reports in unblinded MS trials, will be presented. The 120-day safety update is reviewed, as well as post-marketing data presented in the submission.

Serono has submitted information regarding the orphan drug status of the beta-interferons, and proposed that Rebif should not be blocked from approval at this time by the orphan drug regulations. These issues are not included within the scope of this review document. Reviews discussing these issues have been written by Dr. M. Walton, dated October 1998 and February 1999. These reviews should be examined for all information and assessments with regard to orphan drug issues.

TRIAL ----------

DESIGN

A preliminary version of the protocol for this trial was commented upon by CBER in December, 1993. The final version was not received for comment until after the analysis was completed.

Title: A multicentre, randomised, double-blind, placebo-controlled, phase III study of subcutaneous Rebif™ (recombinant human interferon-beta) in the treatment of relapsing-remitting multiple sclerosis

Code: ----------

Study dates: May 1994 to March 1997

Objectives

The stated primary objective of the trial was to investigate the effects of Rebif at two doses, 6 MIU and 12 MIU, compared to placebo, on the number of exacerbations.

The stated secondary objectives of the trial were to determine the effects of Rebif at the two doses stated above, compared to placebo, on

- duration and severity of exacerbations, time to first exacerbation, and proportion of patients remaining exacerbation-free at 1 and 2 years
- disease activity as measured by numbers of active lesions on monthly cranial T_2-weighted and T_1-weighted Gd-MRI
- burden of disease as measured by cranial T_2-weighted MRI
- deterioration of disability
- safety and tolerability
- need for steroid therapy and hospitalization for MS

Design

This was a double-blinded trial of two dose levels of active agent as compared to placebo, conducted at 22 sites in Canada, and the European Union. Randomization was stratified by center in blocks of 6. Imaging was performed at various sites but analyzed at one site (University of Vancouver, British Columbia). Imaging was performed biannually for all subjects, with two subgroups selected to receive more frequent MRIs. An independent panel performed ongoing review of safety information and supervision of the interim analysis.

Treatment

Trial treatment was to be placebo or Rebif at either 22 mcg or 44 mcg TIW, subcutaneously, for 2 years. It was distributed as 42 vials (a little more than a 3-month supply) containing 0.65 ml of Rebif or placebo solution (the maximal amount to be withdrawn was 0.5 ml). (It should be noted that the formulation for marketing is the same, but is distributed in prefilled syringes.) Placebo was the same solution as active treatment, without interferon; it consisted of sodium acetate, 0.01 M, with mannitol and human serum albumin. Treatments were

to be injected by the subject or a family member at the same time each day, preferably in the evening.

To ameliorate the unblinding effects of sudden administration of full dose of interferon, subjects were to titrate their dose in the initial 8 weeks of the trial. For the first 2-4 weeks of treatment, 1/5 of the volume of trial agent was to be administered; then ½ of the volume for the second 2-4 weeks, followed by the full dose. This schedule was a rough guideline that could be modified individually. Subjects were to be excluded if they could not tolerate doses higher than 1/5 dose by the beginning of the 5th week, or withdrawn if unable to tolerate a full dose by the end of 8 weeks.

Dosing could be adjusted to ½ dose in the event of a persistent WHO grade 2 toxicity; the occurrence of a causally related grade 3 toxicity would allow the reduction or interruption of the dose, with return to full dose or discontinuation related to level of persistent toxicity. Subjects were to be withdrawn from treatment in the event of a grade 4 toxicity due to Rebif. However, there were to be no interruptions or withdrawals for neurological events.

Paracetamol (acetaminophen) was to be given at the discretion of the investigator prophylactically and to ameliorate constitutional symptoms at a dose of 325-1000 mg as required during the treatment period.

Blinding

Techniques designed to maintain the blind included the following:

1) The product was to be sent to sites blinded, with unique subject identification numbers attached to packs and enclosed vials by the manufacturer.

2) Subjects were to be seen by separate treating and evaluating physicians. The treating physician was charged with supervision of trial agent administration, recording and treating of adverse experiences, monitoring of safety assessments including routine laboratory parameters and physical examination, and decisions on the administration of corticosteroids. The evaluating physician was charged with responsibility for all neurological assessments and follow-up of all exacerbations. This evaluating physician was under orders to restrict communications with subjects and other physicians.

3) Subjects were told to cover sites of injections when being seen by evaluating physician and not to discuss any symptoms that might in any way be related to treatment, including injection site reactions.

4) Data from clinical and neurological assessments were to be collected and stored in 2 separate case report form binders

5) Queries about the data from each subject were to be sent to centers as separate sets of clinical queries and neurological queries

6) MRI scans were to be analyzed by technicians whose knowledge of subjects was restricted to subject identification number, date of birth, initials, and whether there was to be another scan.

Subject inclusion criteria

The following were to be fulfilled within 42 days of the first day of treatment:

- age 18-50 years
- clinically definite or laboratory supported RRMS (ref.1) with at least a 1-year history before trial entry
- history of 2 or more exacerbations in the prior 2 years
- stable neurological state for at least 4 weeks at the time of the pre-trial evaluation

- EDSS 0-5.0
- willing and able to comply with the protocol for the duration of the trial
- written informed consent

Subject exclusion criteria

- prior systemic treatment with interferons
- history of lymphoid irradiation
- prior treatment with cyclophosphamide
- prior treatment with azathioprine, cyclosporine A, or any other immunosuppressive in the 12 months prior to trial entry
- prior treatment with corticosteroids or ACTH in the 2 months prior to trial entry
- exacerbation in the prior 3 months (changed to 8 weeks in amendment 1)
- investigational agent or experimental procedure in the prior 12 months
- systemic disease that might interfere with safety, compliance, or evaluation
- history of severe reaction to paracetamol, gentamicin, or gentamicin analogues
- known to have (+) HIV, HTLV-1, and/or hepatitis B serology
- breastfeeding or pregnant or childbearing potential (defined as post-menopausal, surgically sterile, practicing contraception in specified ways, or being sexually inactive)

Concomitant medications

While paracetamol (acetaminophen) was allowed at the discretion of the investigator prior to injections or to treat constitutional symptoms, this therapy was not mandated by the protocol. Nonsteroidal antiinflammatory drugs were only to be used if paracetamol (acetaminophen) failed. Corticosteroids were to be given at the discretion of the treating physician for disabling acute exacerbations (ones defined as interfering with routine daily activities). A standard corticosteroid treatment regime was outlined in the protocol. Immunosuppressives such as azathioprine, cyclosporine A, or others were not allowed to be given during the trial.

Evaluations

Screening: history and physical examination, neurological examination and performance of neurological performance tests, MRI, ECG and CXR, routine hematology, blood chemistry, thyroid function, and urinalysis.

Day 1: (the day of first treatment): complete physical, including neurological examination and performance of neurological performance tests; other tests included serum for baseline antibody measurements, routine hematology, blood chemistry, and urinalysis, and psychological tests.

Treatment period:

- physical examination at the end of months 1,2,3,6,9,12,15,18,21, and 24
- routine hematology, blood chemistry and urinalysis at the end of weeks 2, 4, and 6 and months 2,3,6,9,12,15,18,21, and 24
- thyroid function and serum for antibody testing biannually
- psychological testing (questionnaires) biannually in English-speaking countries only
- MRI:
    - cohort 1 (entire subject pool): biannual PD/T2 imaging

-cohort 2 (subjects at sites 2,5,6,8,10,11,and 13): PD/T2 and Gd-enhanced T1 imaging monthly to month 9
-cohort 3 (subjects at site 5 only): PD/T2 and Gd-enhanced T1 imaging monthly to month 24 (end of trial)

- neurological exam:
  -cohort 1: at the end of months 3, 6, 9, 12, 15, 18, and 24 (on the same day as MRI imaging on appropriate dates, or within 48 hours)
  -cohorts 2 and 3: additional exams to coincide with the monthly MRI imaging
- adverse event recording at each clinic visit

Notes on data collection and analysis

1) Evaluation of exacerbations: severity and duration

Subjects were to inform their center within 48 hours of onset of an exacerbation, with a visit to a site arranged as quickly as possible, preferably within 7 days. The evaluating physician was to determine the nature, severity, and duration of the exacerbation and describe the worst symptoms experienced in each of the functional systems affected. As described in amendment 2, if an exacerbation was brought to the attention of an evaluator during its occurrence, its severity was to have been measured on the Scripps neurological rating scale (which measures functions of the major components of the nervous system); if not evaluated during its occurrence, its severity was to have been determined by its effect on the activities of daily living as interpreted by the Activities of Daily living (ADL) scale. The duration of exacerbations was capped at 3 months for the purposes of analysis in amendment 2.

2) MRI imaging evaluations

Details of the analysis of MR imaging were not in the original protocol. This analysis was performed in a series of steps involving demarcation of regions of interest by radiologists or technicians, followed by outlining of the regions by technicians, then review by radiologists and measurement of areas. An entire data set for an individual subject was analyzed in chronological order by a single technician. A given subject's scans were analyzed as a set by the same technician designated by the study site. No scans were read by more than one technician. Detailed definitions of the different categories of activity on MR were not provided in the protocol.

<u>Primary endpoint</u>

The primary endpoint was the number of protocol-defined exacerbations per subject, using the following definition:

> "..the appearance of a new symptom or worsening of an old symptom, attributable to MS, accompanied by appropriate new neurological abnormality, or focal neurological dysfunction lasting at least 24 hours in the absence of fever, and preceded by stability or improvement for at least 30 days."

<u>Secondary endpoints</u>

Secondary endpoints fell into 3 general categories: exacerbation-related, disability-related, and related to imaging of lesions on MRI. They were not ranked in order of importance by Serono. They were stated as follows:

- disease activity as measured by the number of active lesions on MRI (new, enlarging, recurrent on PD/T2 or enhancing on a T1-Gd scan)
- duration of exacerbations
- severity of exacerbations as defined by changes in the Scripps NRS scale
- time to 1st exacerbation
- proportion of subjects remaining exacerbation-free at 1 and 2 years
- burden of disease as defined by the total area of all lesions on MRI
- deterioration of disability as measured by change in EDSS
- need for steroid therapy and hospitalization for MS

Safety endpoints

Listed safety endpoints included lab tests, vital signs, and psychological status assessed by psychometric testing.

Interim analysis

One interim analysis was planned. It was to occur when 100 subjects/arm had completed 1 year on study. Its purpose was to re-randomize placebo subjects to the treatment groups if efficacy were found in the high-dose group at p=.005.

Serono stated that for analysis of monthly MRI results, the interim analysis would include all the data; thus no adjustment would be considered necessary and the analysis would be final.

Individuals allowed to see the results of the interim analysis were to be the vice president for Medical Affairs, the Therapeutic Director, a responsible biometrician from Serono, and persons not otherwise involved with the trial including an external neurologist, neuroradiologist, biometrician, and a clinical expert in the use of IFN-β.

Final analysis

The original protocol specified the primary analytical population as all subjects reaching the full dose (0.5 ml) of the trial agent by week 8; this was changed in amendment 1 to be an intent-to-treat population defined as all subjects randomized. There was to be no imputation for missing data.

The analytical method for the primary endpoint was to be a Cochran-Mantel-Haenszel ANOVA stratified by center, using observed counts as the scores. All 3 treatments were to be compared simultaneously and pairwise.

The protocol did not specify analytical methods for each secondary endpoint. Rather, a summary statement was made:

> "Categorical variables such as proportions of exacerbation-free subjects or the severity of exacerbation will be analysed by Cochran-Mantel-Haenszel chi-square test, stratified by center. Analysis of variance (ANOVA) taking the center effect into account will be applied for continuous variables. The time to first exacerbation was to be analyzed with the log-rank test, stratified by center. The burden of disease will be measured in terms of percent change because of the variability of the different MRI machines used in the various centres. The changes in EDSS as well as the ambulation index and arm index, will also be analyzed and compared between the treatment groups."

Comments on the analytical plan

The primary endpoint was vaguely written; as an example, the protocol did not define a time point for endpoint analysis of efficacy. The exact measures used for the primary endpoint (mean numbers of exacerbations per subject or count of exacerbations), and for several secondary endpoints, were not specified in the protocol. Definitions of MR lesion categories and statistical analyses of clinical and MRI data were incompletely defined.

RESULTS: CONDUCT OF THE TRIAL

Formal protocol modifications

The protocol was amended 3 times, all after the trial start date in May 1994:

1) Amendment 1 was put into effect August 22, 1994. The major changes were:

- The inclusion criterion, time from the last exacerbation to day 1 of the trial, was reduced from 3 months to 8 weeks.
- A 7-day pretreatment time window for the performance of the baseline neurological examination was added.
- The reporting form for hospitalizations due to MS was changed to the "Notification of Multiple Sclerosis Related Hospitalisation."
- The MRI scanning protocol was revised.
- Further definition of the severity of exacerbations was provided, as follows:

  i. if the worst severity of an exacerbation were experienced during a subject visit or hospitalization, then the event would be scored with specified Scripps Neurological Rating Scale criteria

  ii. if the worst severity of an exacerbation were not to coincide with a neurological assessment then the event would be scored by its effect on the activities of daily living, including the need for hospitalization

2) Amendment 2 was put into effect February 22, 1996. The following changes were made:

- The maximal duration of an exacerbation, for analytical purposes, was set at 3 months
- Extension trial procedures were put in place
- The method of determining the deterioration of disability was defined as the time to deterioration of disability by 1.0 points or a deterioration of 0.5 points between 6.0 and 7.0 on the EDSS, confirmed at 2 consecutive visits 3 months apart

3) Amendment 3 was put into effect on March 5, 1997. It added more details to the performance of an extension trial.

Comment on the amendments

It is unlikely that these protocol changes exerted any bias in the conduct or interpretation of the trial, since they do not seem to affect one treatment group more than another.

Trial enrollment

Five hundred and sixty subjects were randomized and treated; 187 in the placebo group, 189 in the IFN 22 mcg SC TIW group, and 184 in the IFN 44 mcg SC TIW group. The first day of treatment for the first subject was May 11, 1994; for the last, March 2, 1995. Among the 22 sites, 8 enrolled exactly 20 subjects and 6 enrolled exactly 30 subjects. There was not a preponderance of subjects at any one site, with 40 the highest number enrolled at one site

(another site enrolled 39). One site each enrolled 19, 21, 22, 24, 27, and 28 subjects. In addition, subjects were evenly distributed by treatment assignment, with a maximal enrollment disparity among the treatment arms of 2 at only 3 sites.

Comments on enrollment

The trial enrolled considerably more subjects than projected by the initial sample size estimate (around 186 per treatment arm, as opposed to 100 per treatment arm). Serono cites factors such as the rapidity of enrollment and contractual obligations to centers as contributing to this overage. Balance of enrollment, both by site and by treatment group within site, was excellent.

Adherence to schedule of evaluations, followup, and data collection

Baseline neurological evaluations occurred within the protocol-required 7 days of the first day of trial agent administration in all but 4 subjects (at 8 days for two high-dose subjects; 10 days for 1 high-dose subject; and at 14 days for one low-dose subject). Attendance at scheduled visits was excellent: the percent of subjects attending all expected clinical visits (corrected for the amount of time spent in the trial) for the placebo, low-, and high-dose groups 98% (180/184), 95% (179/189), and 96% (180/187).

Table 1 shows the time spent on the study by treatment group.

Table 1. Time on study (days)

	Placebo *n=187*	IFN 22 mcg TIW *n=189*	IFN 44 mcg TIW *n=184*
mean +/- sem	711 +/- 7.3	707 +/- 8.3	723 +/- 5.0
median (Q1, Q2)	730 (729,736)	730 (729, 736)	731 (729,736)

The percentages of subjects followed to the completion of the trial at 24 months was 95.2% overall: the highest number of subjects followed to the completion of the trial was in the high-dose group (94.7% placebo, 93.7% in the low-dose group, and 97.3% in the high-dose group).

MR imaging was very complete: in the placebo, low-, and high-dose groups the percentages of subjects receiving MR scans at the end of the trial were 92% (172/187), 90% (171/189), and 93% (171/184), respectively. Of 4912 PD/T2 scans, only 64 (1.3%) were rejected, of which 40 were redone or replaced. Of the 2767 T1-Gd scans, 12 (0.43%) were rejected due to a protocol violation.

Although the protocol called for the exclusion or withdrawal of subjects in the event of nontoleration at the initiation of treatment (criteria were set at 5 weeks and at 8 weeks), this must have occurred rarely. The number of subjects with times on study of ≤5 weeks in the placebo, low-, and high-dose groups was 0, 1, and 0; the number with times on study of ≤8 weeks was 0, 3, and 0.

Eligibility, dosing, and other violations

Table 2 shows the protocol deviations as tabulated by Serono.

Table 2. Subjects with protocol deviations

type of deviation	placebo *n=187*	IFN 22 mcg TIW *n=189*	IFN 44 mcg TIW *n=184*
age >50 years	0	2	0
concurrent chemotherapy	1	0	0
baseline EDSS>5	1	1	3
hepatitis B +	0	0	1
duration of MS <1 year at trial entry	1	0	2
fewer than 2 exacerbations in last 2 years	2	5	6
exacerbation <8 weeks before trial entry	2	1	0
Interruption of dosing ≥30 days	1	4	10
Less than 100 ml of trial drug	13	19	22
Ended trial early; entered extension trial	0	0	1
Code broken	1	0	0
EDSS missing during trial*	2	7	2
Consent signed after trial procedure (before trial entry)	2	0	4
Totals	26	39	51

*There were 12 instances, with one subject in the low-dose group having 2 missing EDSS evaluations

Interim analysis

The interim analysis took place much later than planned: December, 1996. At this point all but 11 subjects had completed the 2 planned study years. Although the analysis determined benefit at $p<0.005$, it was decided that the study would not be terminated. Serono did not explain the reason for the late performance of the interim analysis.

Comments on the conduct of the trial and data collection

This was a well-conducted trial in which subject eligibility requirements were well fulfilled, the schedule of baseline evaluations was adhered to, followup occurred as planned, and numbers of subjects followed to completion were very good and well balanced among the treatment arms. MRI data collection was excellent.

The bulk of the protocol violations consisted of dosing violations: the administration of less than the protocol-recommended amount of trial agent. These occurred primarily in the active-treated groups, in a dose-related manner. The number of interruptions of ≥60 days, a "major" violation, in the placebo, low- and high-dose groups was 0, 2, and 4, respectively. There were also more discontinuations of treatment in the active-treated groups (see the Safety section of this review). While these violations may reflect the tolerability of IFN, they would not affect the efficacy conclusions of the trial.

The protocol amendments affecting the conduct and analysis of the trial do not appear to have a biasing effect of one treatment over another.

RESULTS: BASELINE CHARACTERISTICS OF THE SUBJECTS

Table 3 shows important baseline subject characteristics.

Table 3. Selected baseline subject data

	placebo *n=187*[1]	IFN 22 mcg TIW *n=189*[1]	IFN 44 mcg TIW *n=184*[1]
% male/ female	25/75	33/67	34/66
age (years)*	34.7	34.8	35.2
height (cm)*	169.9	170.1	170.0
weight (kg)*	68.0	69.2	70.2
% Caucasian/Negroid/other	98/0/2	99/0/1	99/0.5/0.5
% clinically definite/laboratory defined	86/14	87/13	89/11
time since onset (years)	6.1	7.7	7.8
number of exacerbations within 2 years of trial start*	3.0	3.0	3.0
number of treated exacerbations within 2 years of trial start*	1.5	1.3	1.3
% previously receiving corticosteroids or ACTH	82	76	82
baseline EDSS*	2.4	2.5	2.5
MRI lesion total area on PD/T2 scans (mm^2)	3082	2818	2975

*Means are shown. Medians, standard deviations, minima and maxima as well as 1st and 3rd quartiles were similar among treatment groups
[1]except for weight: placebo, n=185; IFN 22 mcg TIW, n=188; IFN 44 mcg TIW, n=183

Comments on baseline characteristics of subjects

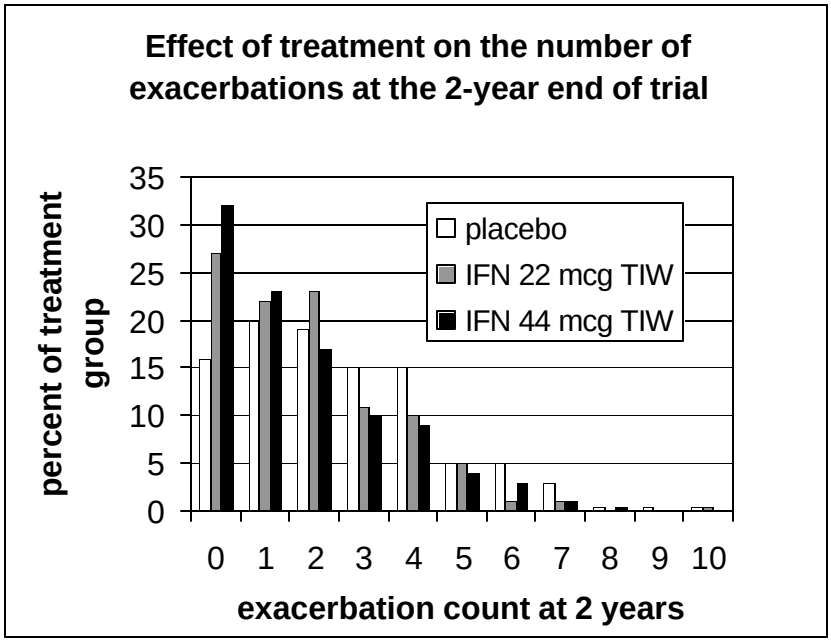
Baseline characteristics of subjects were well balanced among the treatment arms. Race distribution was characteristic of the MS population. The ratio of males to females more closely approximated the overall population ratio (1:2) in the two treatment groups than in the placebo group, where the ratio of males was lower (1:3). There was a slightly lower time since the onset of disease in the placebo group, but indices of disease activity and extent were similar.

RESULTS: EFFICACY

Primary endpoint: Numbers of exacerbations/subject

Figure 1 illustrates the percent of each treatment group with exacerbations at 2 years. Treatment tended to yield larger numbers of subjects with fewer exacerbations. Note that Serono attributed 0 exacerbations to those subjects who were lost to followup without an exacerbation, such subjects were found only in the placebo and low-dose groups. This manipulation was performed for 2 and 6 subjects respectively at 2 years (1.1 and 3.2% of subjects at 2 years).

Figure 1. Distribution of numbers of exacerbations at the 2-year end of the trial



The distribution of exacerbations at 1 year followed the same pattern; the maximal number of exacerbations for that period was 5. Subjects lost to followup: 1 and 4 subjects in the placebo and low-dose groups at 1 year (0.5% and 2.1% of subjects).

Table 4 shows Serono's statistical analysis of the primary endpoint: mean numbers of exacerbations experienced by subjects in each treatment arm.

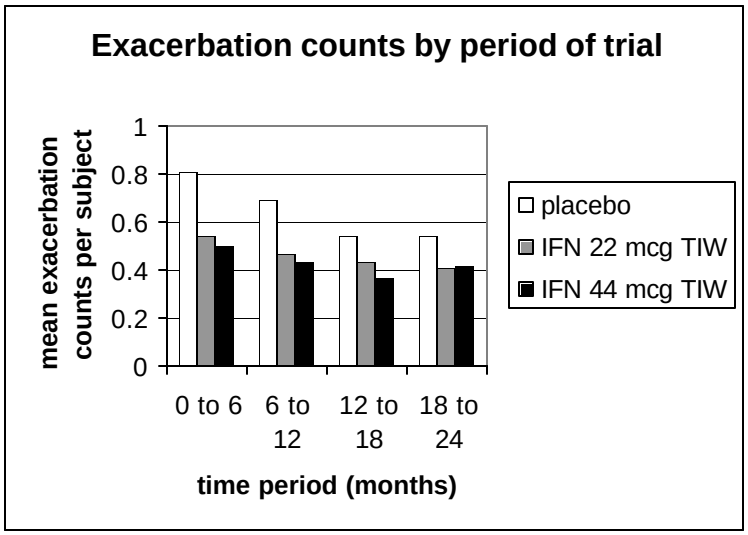
Table 4. Mean exacerbation count per subject at 1 and 2 years

time in study	placebo *n=187*	IFN 22 mcg TIW *n=189*	IFN 44 mcg TIW *n=184*	p-value IFN 22 vs placebo	p-value IFN 44 vs placebo	p-value IFN 22 vs IFN 44
1 year	1.5	1.0	0.9	<0.0001	<0.0001	0.34
2 years	2.56	1.82	1.73	0.0002	<0.0001	0.37

Note: statistical comparisons were performed using a log linear model taking into account center and time on study. The numbers of subjects with a <95% of time on study in placebo, low, and high-dose groups were 11, 12, and 5, respectively.

Figure 2 illustrates data provided by Serono on exacerbation counts over time.

Figure 2. Exacerbations counts (means) corrected for subject years in a given period.



Numbers of subjects

time period (months)	placebo	IFN 22 mcg TIW	IFN 44 mcg TIW
0 to 6	187	189	184
6 to 12	184	186	183
12 to 18	182	182	182
18 to 24	177	178	180

Rates of exacerbations dropped more in the placebo group than in the treatment groups over time, but the treatment effect persisted. At no time was there a remarkable difference between the two treatment groups.

Upon request, Serono provided a statistical analysis of exacerbations in the 2nd year of the trial. The analysis method was the same as that provided in the original application for 1- and 2-year exacerbation counts. The difference between either treatment group and placebo remained statistically significant during the second year of the trial.

Table 5. Mean exacerbations per subject in the 2nd year of the trial only

Placebo n=182	IFN 22 mcg TIW *n=182*	IFN 44 mcg TIW *n=182*	p-value IFN 22 vs placebo	p-value IFN 44 vs placebo	p-value IFN 22 vs IFN 44
1.09	0.85	0.82	0.03	0.007	0.059

Note: the analytical procedure was a log-linear model taking into account center and time on study

Comment

A large percentage of subjects in the active groups suspected that they were on active treatment (see Appendix: blinding questionnaire, Table 33). Since the primary endpoint depended to some extent upon reporting by patients this raises the concern that the endpoint was subject to some bias.

CBER sensitivity analyses of primary endpoint

1) Presented analysis (log-linear model) compared to planned analysis (Cochran-Mantel-Haenszel ANOVA)

The analysis presented above of the primary endpoint was different from the analysis specified in the protocol, the Cochran-Mantel- Haenszel ANOVA. Both analyses take into account the effect of center on exacerbation counts; the presented analysis further takes into account the subjects' times on study. The analyses reached the same overall conclusion regarding efficacy; for the comparisons of each active treatment to placebo at 1 or 2 years, the p-values were 0.001, and for the comparisons of low-dose to high-dose, the p-values at 1 and 2 years were 0.36 and 0.49, respectively. The results were further confirmed by CBER's unadjusted (for the center and subjects' time on study) and logistic regression analyses. The overall conclusions did not change. Thus Serono's use of a non-protocol-defined analytical technique did not change the overall conclusions from the trial.

2) Exacerbations potentially not counted based on proximity of onset to previously occurring ones

The calculation of exacerbations was based on information collected on an exacerbation report form. Recorded exacerbations were assumed to fit the protocol definition (see "endpoints and primary analyses, primary endpoints), except that those occurring within 1 month of each other were considered as 1 exacerbation. Examination of line listings showed that this time-related adjustment was performed for 11 placebo, 11 low-dose, and 2 high-dose subjects, resulting in a loss to analysis of 16/478 (3%), 14/344 (4%), and 3/319 (1%) potential exacerbations respectively over the 2 years of the trial. Additionally, the numbers of these "not-counted" exacerbations for which corticosteroids were administered was 6, 5, and 0, respectively; other, symptomatic treatments were not recorded for exacerbations generally. Serono's analysis thus possibly undercounted the number of exacerbations and corticosteroid-treated exacerbations for placebo and low-dose subjects more than for the high-dose. This analysis would not change a conclusion of efficacy regarding the comparison of placebo to high dose, and would likely have little effect on the comparison of placebo to low dose.

3) Imputations for missing observations

Serono's analysis of the numbers of exacerbations attributed 0 exacerbations to 5 subjects at 1 year (1 placebo, 4 low-dose) and 8 subjects at 2 years (2 placebo, 6 low-dose) with missing observations. CBER determined p-values (unadjusted for time on study), using a log-link procedure, deleting the missing values and attributing the median and worst values overall to the missing ones. As can be seen in Table 6, none of the analyses changed the conclusions regarding the primary endpoint as presented by Serono.

Table 6. P-values for the primary endpoint, mean numbers of exacerbations (unadjusted for time on study), using different imputation methods

	placebo vs. high dose	placebo vs. low dose	low vs. high dose
deletion	0.0001	0.0005	0.21
median	0.0001	0.0005	0.17
worst	0.0001	0.0027	0.08

4) Site and treatment effect

Table 7 shows mean exacerbation counts by treatment group and center. The effect of treatment was consistent, i.e., with few exceptions, each dose resulted in fewer mean numbers of exacerbations at 2 years than placebo.

Table 7. Mean 2-year exacerbation counts by center

Site code	placebo	*n*	IFN 22 mcg TIW	*n*	IFN 44 mcg TIW	*n*
1	3.00	*7*	1.17	*6*	1.57	*7*
2	1.70	*10*	2.50	*10*	1.60	*10*
3	3.15	*13*	1.36	*14*	1.92	*13*
4	1.90	*10*	1.80	*10*	0.60	*10*
5	1.92	*13*	2.36	*14*	2.25	*12*
6	2.4	*10*	1.90	*10*	1.30	*10*
8	3.00	*10*	2.50	*10*	2.20	*10*
9	1.67	*6*	0.71	*7*	1.33	*6*
10	5.14	*7*	4.67	*6*	5.00	*8*
11	3.6	*10*	2.25	*8*	2.33	*9*
12	2.57	*7*	2.57	*7*	2.50	*6*
13	3.00	*10*	1.56	*9*	1.33	*9*
14	1.71	*7*	1.50	*8*	0.43	*7*
15	1.57	*7*	2.00	*6*	0.29	*7*
17	2.60	*10*	1.50	*10*	0.80	*10*
18	2.00	*8*	0.75	*8*	1.00	*8*
19	3.40	*10*	2.00	*10*	2.60	*10*
21	2.50	*6*	1.50	*8*	2.67	*6*
22	2.43	*7*	1.00	*7*	1.50	*6*
23	2.29	*7*	1.14	*7*	2.50	*6*
26	2.50	*6*	2.14	*7*	1.57	*7*
27	1.50	*6*	1.14	*7*	0.71	*7*
totals	2.56	*187*	1.82	*189*	1.73	*184*

Comments regarding robustness of primary endpoint

Conclusions regarding mean exacerbation count, the primary endpoint, did not depend upon the method of analysis (Cochran-Mantel-Haenszel vs. log-linear method) nor to methods of imputation of missing observations. Exploration of noncounted exacerbations due to proximity to previous ones suggested no bias against the placebo arm using this procedure. Conclusions regarding the primary endpoint did not appear to be driven by a preponderance of effect at selected sites. The primary endpoint was robust to these analyses.

Secondary endpoint: Duration of exacerbations

Serono found no remarkable or statistically significant effect of either treatment on the duration of exacerbations, with a mean of 47-48 days for all groups.

Comment

The protocol did not specify criteria for determining the duration of exacerbations. Thus this determination was subject to considerable individual investigator variability, possibly rendering detection of differences difficult.

Secondary endpoint: Exacerbation severity

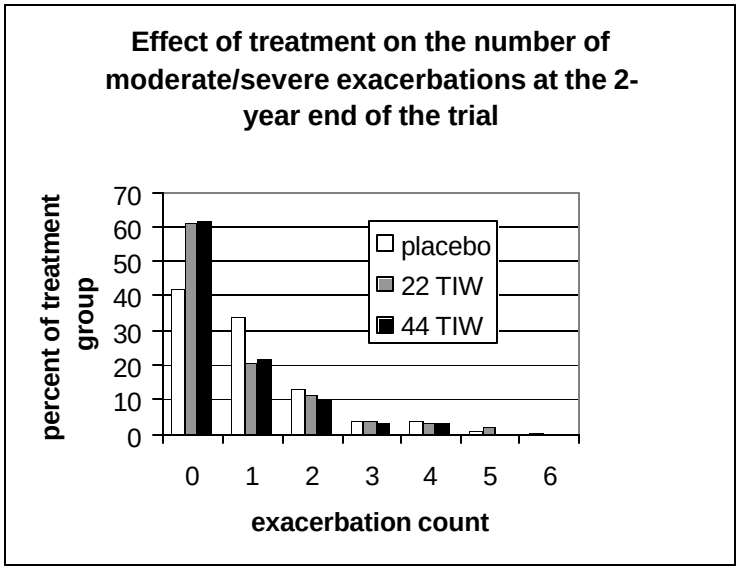
Table 8 shows Serono's analysis of the mean numbers of exacerbations per subject by severity. The distribution of the numbers of moderate or severe exacerbations by treatment group paralleled that for exacerbations of all degrees of severity as presented above.

Table 8. Mean exacerbation count per subject by severity

severity of exacerbation	placebo *n=187*	IFN 22 mcg TIW *n=189*	IFN 44 mcg TIW *n=184*
mild	1.56	1.11	1.11
moderate	0.79	0.56	0.49
severe	0.20	0.15	0.13
any	2.56	1.82	1.73

Figure 3 shows the distribution of subjects with moderate-severe exacerbations by treatment group at 2 years. It shows that there were more subjects with the least numbers of moderate-severe exacerbations in the active treatment groups.

Figure 3. Effect of treatment on the number of moderate-to-severe exacerbations



Serono's statistical test, ANOVA on the ranks taking center into account, showed significance for both active treatments compared to control (p=0.0003 and 0.003 for the high and low-doses compared to control), with no difference between treatments (p=0.53). Serono also did a statistical analysis (ANOVA on ranks taking center into account) excluding those who did not experience an exacerbation. This analysis showed more marginal significance: p-values for the high-and low-dose groups against control were 0.06 and 0.09, respectively, with a value of 0.81 for the intertreatment difference.

Secondary endpoint: Time to 1st (and 2nd) exacerbation

Table 9 shows the time to 1st and 2nd exacerbation as presented by Serono. The time to 2nd exacerbation was not prospectively defined as an endpoint.

Table 9. Median time to 1st and 2nd exacerbation (days)

	Placebo *n=187*	IFN 22 mcg TIW *n=189*	IFN 44 mcg TIW *n=184*	p-value IFN 22 vs placebo	p-value IFN 44 vs placebo	p-value IFN 22 vs IFN 44
1st exacerbation	135	229	288	0.0008	<0.0001	0.16
2nd exacerbation	449	702	Not reached	0.002	<0.0001	0.12

Note: the statistical test was a Cox proportional hazards model taking center into account

Secondary endpoint: Proportion of subjects exacerbation-free

Table 10 shows the proportion of subjects exacerbation-free, as presented by Serono. Subjects who were lost to followup without an exacerbation were censored from the analysis.

Table 10. Exacerbation-free subjects (% of group)

time in study	placebo	IFN 22 mcg TIW	IFN 44 mcg TIW	p-value IFN 22 vs. placebo	p-value IFN 44 vs. placebo	p-value IFN 22 vs. IFN 44
	% *(n)*	% *(n)*	% *(n)*			
1 year	22 *(186)*	37 *(185)*	45 *(184)*	0.0009	<0.0001	0.11
2 years	15 *(185)*	25 *(183)*	32 *(184)*	0.014	<0.0001	0.08

Note: the statistical test was a logistic regression taking center into account.

CBER explorations of selected secondary endpoints related to exacerbations

1) Exacerbation durations: analytical manipulations

Maximal exacerbation duration was defined in amendment 2 as follows: "Any exacerbation that has not resolved either partially or completely after more than 3 months from the date of onset would be defined as 3 months in duration." However, duration was calculated only if the exacerbation occurred more than 3 months before the last subject visit; otherwise, the duration was deemed "missing." In addition, Serono attributed a duration of 90 days to exacerbations whose end date was missing if they occurred within the time window for all exacerbations.

The truncation of the defined duration of an exacerbation for lack of resolution at 3 months was performed in equal percents of each group (72/478(15%), 55/344 (16%), and 54/319 (17%) of exacerbations in the placebo, low, and high-dose groups respectively). Similarly, the exclusion from analysis for late exacerbations was performed similar numbers of placebo, low-, and high-dose cases (46, 39, and 42 cases). In addition, although not specified in the protocol, Serono attributed a duration of 90 days to exacerbations whose end date was missing if they occurred within the time window for all exacerbations. This occurred for similar numbers of cases in placebo, low-, and high-dose groups (20,15, and 20 exacerbations,

respectively). Thus there was no bias seen in various manipulations of the duration of exacerbations.

2) Severity of exacerbations: different scoring systems

To ascertain whether the scoring system used had an effect on the distribution of severity scores among the treatment groups, Serono was asked to present the percentages of exacerbation severities encoded by either the Scripps NRS, the ADL, or both. These percentages for each scale were similar for the treatment groups (about 50% for the Scripps NRS, about 45% for the ADL, the rest being scored by both tests). For an additional request, Serono repeated the analyses of moderate and severe exacerbations as presented above, using data from each scale in turn. This analysis was consistent with the analysis of all moderate and severe exacerbations as presented above. These analyses support Serono's conclusions using the two scoring systems in this trial.

Comments regarding secondary, exacerbation-related endpoints

The effects of IFN were consistent across most of the exacerbation-related secondary endpoints: presence or absence of exacerbation and time to exacerbations and their severity. One endpoint, duration, was not affected by IFN treatment. The effect of IFN treatment on the incidence of exacerbations of each severity paralleled that for overall exacerbations. Thus it is not clear that there was an independent effect on severity of exacerbations.

Secondary endpoint: Deterioration in disability (time to confirmed disability and percent progressors)

Figure 4 shows the Kaplan-Meier analysis of the time to first 3-month confirmed progression in EDSS disability, censoring those lost to followup without a progression. The definition of confirmed deterioration in disability was established in amendment 2 to the protocol.

Figure 4. Percents of subjects with 3-month confirmed EDSS deterioration

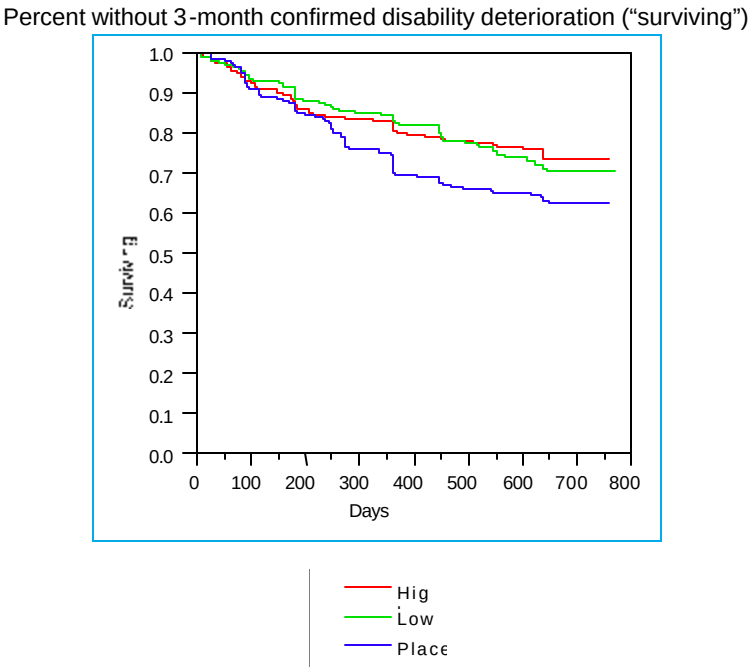


Table 11 shows Serono's statistical analysis of the time to progression in disability. The p-values presented are for a test comparing the 1st quartile of the time to 3-month confirmed progression, not for the entirety of the data regarding time to disability.

Table 11. Confirmed progression in disability

	Placebo *n=189*	IFN 22 mcg TIW *n=189*	IFN 44 mcg TIW *n=184*	p-value IFN 22 vs. placebo	p-value IFN 44 vs. placebo	p-value IFN 22 vs. IFN 44
First quartile of time to disability (months)	11.9	18.5	21.3	0.04	0.014	0.65
Percent (n) progressing*	38 (*69/180*)	30 (*54/182*)	27 (*48/179*)	Not presented		

Note: the statistical test used the Cox proportional hazards model taking center into account, censoring those lost to followup without a progression.
*Excludes those lost to followup without a progression: 7 placebo, 7 low-dose, and 5 high-dose

CBER sensitivity analyses on disability

Because deterioration in disability is of major clinical concern, CBER performed additional analyses of the disability endpoint.

1) Alternative analysis of time to confirmed progression in disability

CBER performed statistical analysis on the entirety of the data on confirmed progression in disability (not only on the first quartile of time to confirmed progression). Table 12 shows the results.

Table 12. Kaplan-Meier analysis of 3-month confirmed deterioration in disability

	placebo	IFN 22 mcg TIW	IFN 44 mcg TIW	IFN 22 mcg vs. IFN 44 mcg
p-value vs. placebo*	-	0.07	0.03	0.63
% of group with 3-month confirmed deterioration at 2 years	37.7	29.4	26.4	-

*statistical test is log-rank

These results demonstrate more marginal levels of statistical significance than the results presented by Serono, but are basically confirmatory of the effects of IFN on increasing the time to progression in disability. As Serono concluded, there was no statistically significant difference between the two doses on prolongation of time to 3-month EDSS confirmed disability.

2) End-of-trial confirmed deterioration in disability status

The time to confirmed progression in disability analysis measures the first time that a confirmed deterioration occurs in a subject's status. However, in some cases status may improve. Subjects may end the trial unchanged or even better than when they started. CBER determined the percentages of subjects who ended the trial with a 3-month confirmed deterioration in disability. Two analyses were performed (Table 13):

- analysis 1: subjects with both month 21 and month 24 EDSS values
- analysis 2: similar to analysis 1, with addition of subjects a) whose last values were before the end of the trial, where the last value represented a deterioration and the previous value also represented a deterioration (these values may have been more than 3 months apart) and b) those with a month 24 value, but whose previous value was more than 3 months prior.

Table 13. Subjects with a 3-month confirmed disability deterioration at the end of the trial

	placebo	IFN 22 mcg TIW	IFN 44 mcg TIW
	% (n)	*% (n)*	*% (n)*
analysis 1 (see text)	14.7 *(177)*	10.9 *(174)*	9 *(178)*
analysis 2 (see text)	15.2 *(184)*	11.8 *(186)*	8.7 *(183)*

Active treatment tended to reduce the numbers of subjects with confirmed deterioration in disability at the end of the trial, with a trend toward a dose effect. Note that the percentages of subjects ending the trial with a confirmed progression are consistently lower (less than a half) than the percentages exhibiting a 3-month "confirmed progression" at any time during the trial.

This suggests that the 3-mopnth confirmation, as used in this trial, is not a good criterion for permanent progression of disability.

Comments on the disability endpoint

Statistical analysis of the time to confirmed deterioration in disability using the entirety of the data over the trial was slightly different from the analysis using the first quartile of deterioration, as provided by Serono. Serono's choice of analysis for p-value is highly unusual for this type of analysis. The log-rank on the full data set is the preferred method. However, the conclusions are basically the same: both doses of Rebif prolonged the time to confirmed deterioration in EDSS disability. The numbers of subjects experiencing a lasting deterioration was small overall; however, there was a trend toward improvement in end-of-trial disability.

Secondary endpoints: MRI

Secondary endpoint: MRI PD/T2 lesion area

Table 14 shows the baseline PD/T2 area for the three groups and the percents change. Baseline values were similar but showed a wide range. An increase in area is regarded as a worsening.

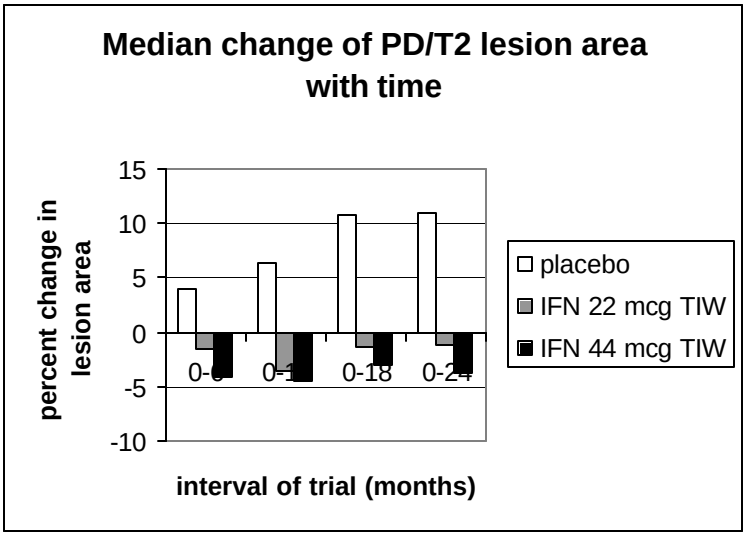
Table 14. MRI PD/T2 area: baseline (mm^2) and percent change at 24 months

		placebo	IFN 22 mcg TIW	IFN 44 mcg TIW
baseline	*n*	*187*	*189*	*184*
	mean	3082	2818	2975
	median	2099	1963	1903
	range	40, 18302	18, 17269	31, 13774
24-month % change	*n*	*172*	*171*	*171*
	mean	24	19	-0.5
	median	11	-1.2	-3.8
	range	-42, 559	-49, 1728	-71, 118

Using ANOVA on the ranks of 24-month change in lesion area, adjusting for center and baseline PD/T2 lesion area, Serono found that the difference between placebo and low- and high-dose IFN was <0.0001 in each case; the p-value of the difference between the two treatments was 0.05.

Figure 5 shows the effect of IFN treatment on median percent change in lesion area.

Figure 5. Median percent change in MRI PD/T2 lesion area with time



Numbers of subjects

Time period (mths)	placebo	IFN 22 mcg TIW	IFN 44 mcg TIW
0-6	182	182	182
0-12	179	180	180
0-18	176	177	172
0-24	172	171	171

The PD/T2 lesion area decrease associated with IFN occurred at the earliest time point examined (6 months) and remained constant over the duration of the trial. The median percent change in PD/T2 lesion area in the placebo group increased to the 18 month time point, where it plateaued.

Comments

Placebo-subject PD/T2 lesion area worsened during the trial. There was a negligible effect on mean change from baseline over the trial in the low-dose group (not shown), but the medians show that there was a shift to smaller lesion areas in this group. High-dose IFN effected a stabilization of lesion areas, starting early in the trial.

Secondary endpoint: MRI "disease activity"

Definitions

"Disease activity" was defined 3 ways by Serono. These were not well-defined in the protocol. In the analysis of new activity, unique identifiers were applied to lesions in an effort to avoid double-counting.

- "Activity" on PD/T2 MRI scans was determined as new, enlarging, persistently enlarging, and recurrent lesions. PD/T2 "activity" was determined for all subjects on their biannual scans, excluding the baseline scan for those not in cohorts 2 or 3 (cohorts 2 and 3 had a prestudy scan, allowing the determination of activity at baseline).
- "Activity" on T1-weighted Gd-enhanced scans was determined as enhancement by the contrast agent, gadolinium, only administered to cohorts 2 and 3. For these analyses, new

activity could be assessed at baseline, as these subjects had a prestudy scan in addition to their baseline scan.

- Combined (T1 +T2) activity was defined as lesions on a scan that had characteristics of T2 and T1 active lesions; this was determined for each of two subgroups (cohorts 2 and 3) in the same way. Serono used the sum of unique, newly active lesions and unique, recurring active lesions.

Cohort 2 consisted of 198 subjects; 7 subjects did not meet the inclusion criteria. Cohort 3 consisted of the 39 subjects at center 5.

Table 15 is a summary of results reported by Serono for various measures of disease "activity" as measured by MRI changes. The percentages of active scans were not specified secondary endpoints.

Table 15. Summary table of MRI "activity" measures (medians*)

		Placebo	IFN 22 mcg TIW	IFN 44 mcg TIW	p-value IFN 22 vs placebo	p-value IFN 44 vs placebo	p-value IFN 22 vs IFN 44
PD/T2 "activity" (all subjects)	*number in cohort*	*184*	*185*	*182*			
	lesions/ subject/scan	2.25	0.75	0.5	<0.0001	<0.0001	0.0003
	% active scans /subject	75	50	25	<0.0001	<0.0001	0.0002
combined (T2+T1) "activity" (cohort 2)	*number in cohort*	*66*	*64*	*68*			
	lesions/subject/ scan	0.88	0.17	0.11	<0.0001	<0.0001	0.16
	% active scans/ subject	44	13	11	<0.0001	<0.0001	0.12
combined (T2+T1) "activity" (cohort 3)	*number in cohort*	*13*	*14*	*12*			
	lesions/subject/ scan	0.9	0.1	0.02	0.09	0.01	0.26
	% active scans/ subject	52	11	2	0.09	0.01	0.28

The statistical test was ANOVA on ranks, adjusted for baseline counts and center (adjustment for center not performed for cohort 3, which was only at center 5).
*Differences among the means were of less relative magnitude in all cases, but trended in the same directions for all of the measures

IFN treatment reduced both numbers of lesions per scan and the numbers of active scans; both the low- and high-dose treatment were significantly different from placebo in measures involving PD/T2 lesions for the whole subject population, and for combined active lesions for the subgroups (cohorts 2 and 3). The difference between high- and low-dose treatment was

significant for the PD/T2 activity measure (whole subject population) but not for combined activity (PD/T2 and Gd-enhanced lesions) in the subgroups.

Comments on MR imaging

The MR results corroborate the clinical effects of IFN treatment. IFN treatment reduced MR measures of MS pathology and activity, with effects that occurred early. It is of interest that there were statistical differences between the two doses of IFN in certain MR imaging results, while the clinical effects tended not to show statistical differences between the two doses. This suggests that MR may be more sensitive to the effects of IFN. This result would not be surprising, since MR may show MS pathology in the absence of clinical manifestations.

Dr. Robert Grossman of the University of Pennsylvania Department of Radiology reviewed MRI technique and analysis in trial ----------. He concludes that this was a "carefully performed study which appears to demonstrate significant differences between treated MS and control patients." He judged the clinically significant results to be: "..(1) The ability of the drug to decrease lesion burden as defined by T2-weighted MR lesions over the duration of the study. (2) The ability to decrease lesion activity as defined by the sum of T1 enhanced lesion numbers and changes in size of additional non-enhancing T2-weighted lesions." However, he expressed a concern about "the inability of independent observers to verify any data based on the large interobserver variability of the methodology," and recommended that a review of the readings to assess the effect of this variability be performed.

<u>Secondary endpoints: Hospitalizations and steroid treatments for MS</u>

Table 16 shows Serono's analysis of hospitalizations for any MS symptom.

Table 16. Hospitalizations for any MS symptom

	placebo *n=187*	IFN 22mcg TIW *n=189*	IFN 44 mcg TIW *n=184*
number of hospitalizations: mean (minimum, maximum)	0.48 (0, 7)	0.38 (0, 4)	0.25 (0,3)
% (number) of subjects hospitalized at least once*	25% (*44/179*)	25% (*44/179*)	18% (*33/181*)

* Subjects lost to follow-up without hospitalizations excluded (8 placebo, 10 low-dose, and 3 high-dose).

Hospitalization rate was defined as the number of hospitalizations divided by time on study. ANOVA on the ranks taking center into account showed that the p-value for the difference from placebo for this parameter was 0.038 for the high-dose and 0.57 for the low-dose, with no statistical difference between the two active groups.

Table 17 shows steroid courses for MS symptoms.

Table 17. Steroid courses for MS symptoms

	placebo *n=187*	IFN 22mcg TIW *n=189*	IFN 44 mcg TIW *n=184*
number of all steroid treatments for MS: mean (minimum, maximum)	1.39 (0, 11)	0.97 (0, 8)	0.75 (0, 7)
% (number) of subjects, steroids used at least once*	57% (*104/181*)	44% (*80/182*)	39% (*71/183*)

* Subjects lost to follow-up without a steroid treatment excluded (6 placebo, 7 low-dose, and 1 high-dose).

Steroid use rate was defined as the number of steroid treatments divided by time on study. ANOVA on the ranks taking center into account showed that the p-value for the difference from placebo was 0.0002 for the high-dose and 0.013 for the low-dose, with no statistical difference between the two active groups.

Comments on hospitalizations and MS-related use of steroids

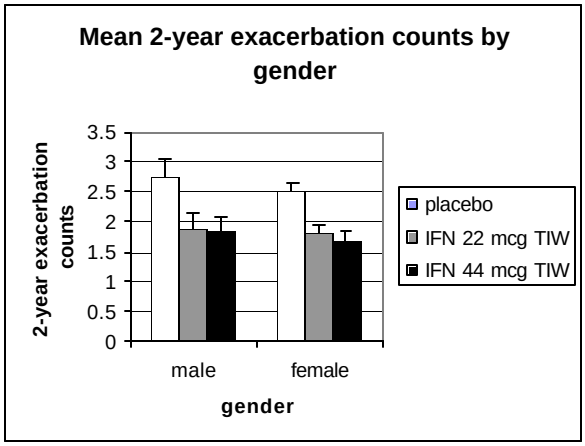
The analyses of hospitalization rates and steroid courses for MS support the overall conclusion of the effect of active treatment with IFN on exacerbations.

CBER exploratory analyses for differential efficacy

CBER examined exacerbation counts, deterioration in disability, and change in MRI lesion area in the subgroups of gender and baseline age, weight, EDSS, and MRI lesion area in an effort to ascertain if there is compelling evidence of a lack of benefit in any of these important categories of subjects. Figures 6 through 10 illustrate these analyses. Subjects with missing data were omitted.

Figure 6. Efficacy and imaging outcomes by gender (means shown with standard errors).

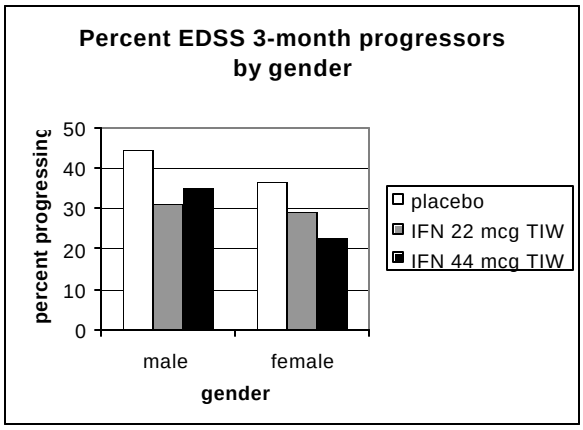
a) 2-year exacerbation counts



Numbers of subjects

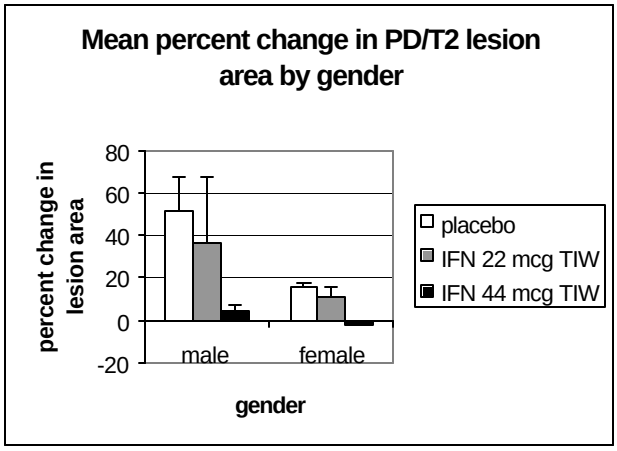
gender	placebo	IFN 22 mcg TIW	IFN 44 mcg TIW
male	46	63	62
female	141	126	122

b) EDSS 3-month confirmed progressors



Numbers of subjects

gender	placebo	IFN 22 mcg TIW	IFN 44 mcg TIW
male	45	61	60
female	135	121	119



c) Mean percent change in MRI PD/T2 lesion area

Numbers of subjects

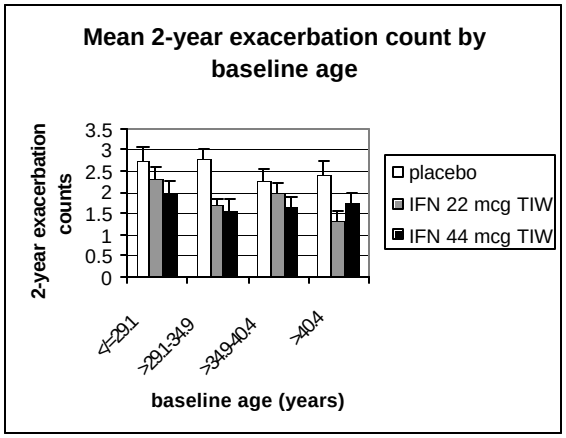
gender	placebo	IFN 22 mcg TIW	IFN 44 mcg TIW
male	41	55	58
female	131	116	113

Comment

Examination of the placebo groups shows that the course of MS during the trial was worse in males than in females. However, Rebif at both doses resulted in benefit for both sexes, with no clear trend toward a difference in degree of benefit.

Figure 7. Efficacy and imaging outcomes by baseline age (means shown with standard errors).

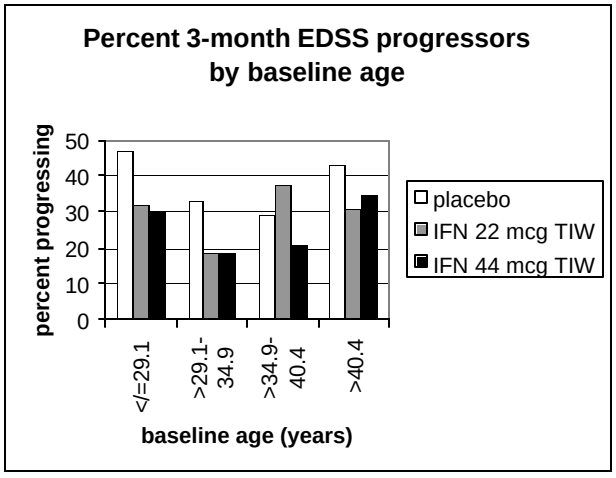
a) 2-year exacerbation counts



Numbers of subjects

baseline age (yrs)	placebo	IFN 22 mcg TIW	IFN 44 mcg TIW
≤29.1	47	44	48
>29.1-34.9	53	51	39
>34.9-40.4	40	54	44
>40.4	47	40	53

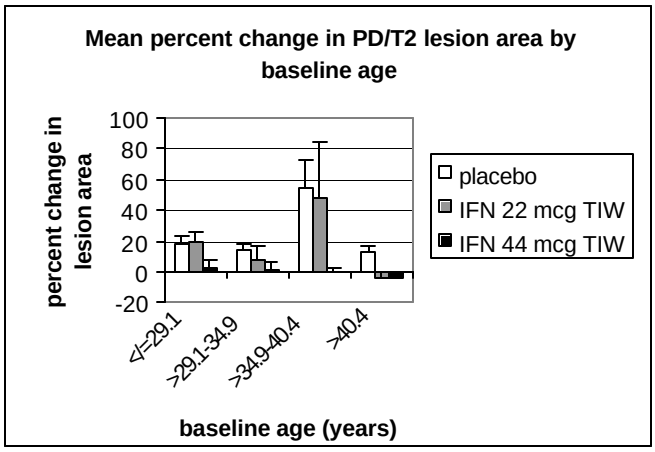
b) EDSS 3-month confirmed progressors



Numbers of subjects

baseline age (yrs)	placebo	IFN 22 mcg TIW	IFN 44 mcg TIW
≤29.1	47	41	46
>29.1-34.9	51	49	38
>34.9-40.4	38	53	43
>40.4	44	39	52

c) Mean percent change in MRI PD/T2 lesion area



Numbers of subjects

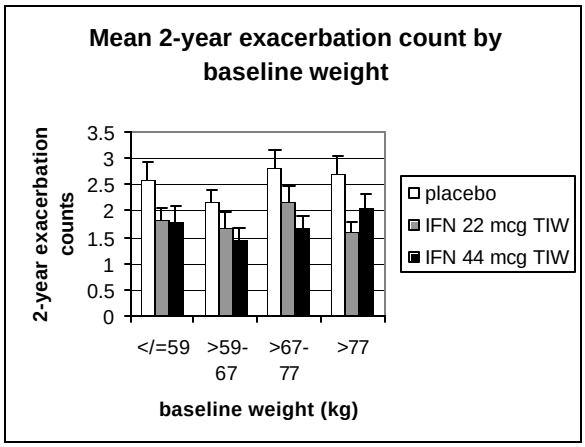
baseline age (yrs)	placebo	IFN 22 mcg TIW	IFN 44 mcg TIW
≥29.1	45	37	43
>29.1-34.9	48	47	35
>34.9-40.4	37	48	42
>40.4	42	39	51

Comment

With some sporadic exceptions, both doses of IFN effected benefit for all age groups, with no clear trend toward less benefit at either extreme of age.

Figure 8. Efficacy and imaging outcomes by baseline weight (means shown with standard errors).

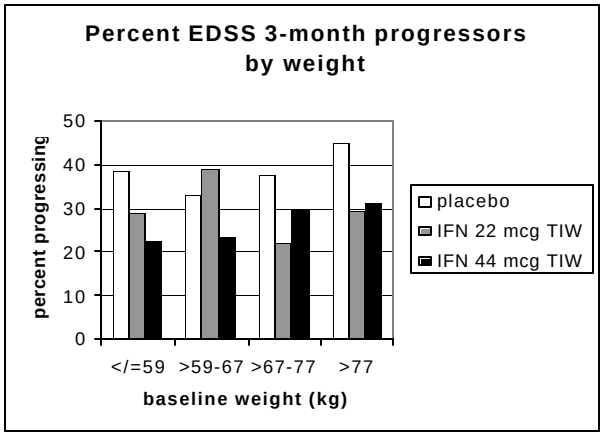
a) 2-year exacerbation counts



Numbers of subjects

baseline weight (kg)	placebo	IFN 22 mcg TIW	IFN 44 mcg TIW
≤59	54	55	41
>59-67	50	44	47
>67-77	43	42	48
>77	40	48	48

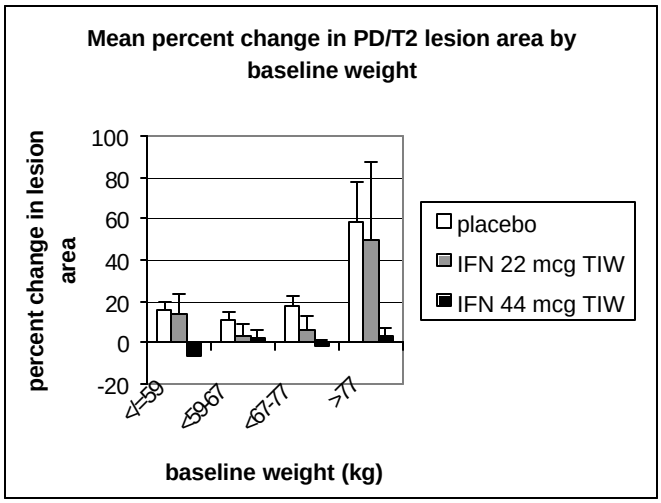
b) EDSS 3-month confirmed progressors



Numbers of subjects

baseline weight (kg)	placebo	IFN 22 mcg TIW	IFN 44 mcg TIW
≤59	52	52	40
>59-67	48	41	47
>67-77	40	41	47
>77	40	48	45

c) Mean percent change in MRI PD/T2 lesion area



Numbers of subjects

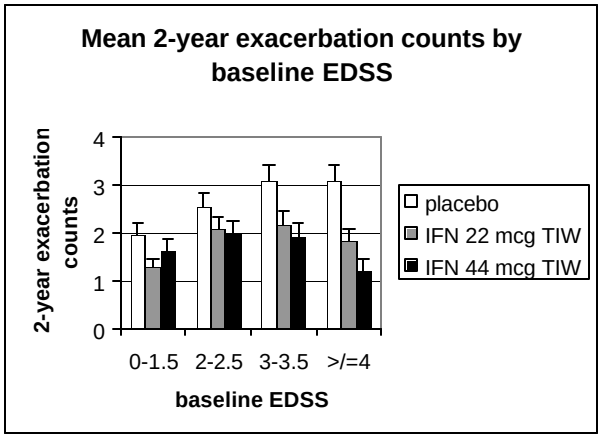
Baseline weight (kg)	placebo	IFN 22 mcg TIW	IFN 44 mcg TIW
≤59	51	46	39
>59-67	47	41	43
>67-77	38	38	44
>77	36	46	45

Comment

With one sporadic exception both doses of IFN effected benefit for all weight groups, with no clear trend toward less benefit at either extreme of weight.

Figure 9. Efficacy and imaging outcomes by baseline EDSS (means shown with standard errors).

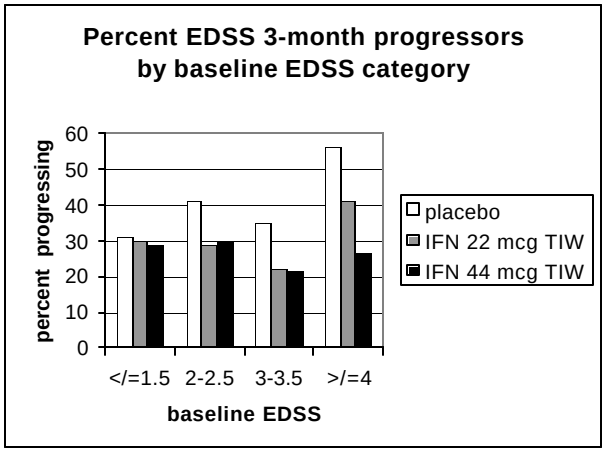
a) 2-year exacerbation counts



Numbers of subjects

baseline EDSS	placebo	IFN 22 mcg TIW	IFN 44 mcg TIW
≤1.5	62	57	58
2-2.5	53	51	48
3-3.5	44	46	47
≥4	28	35	31

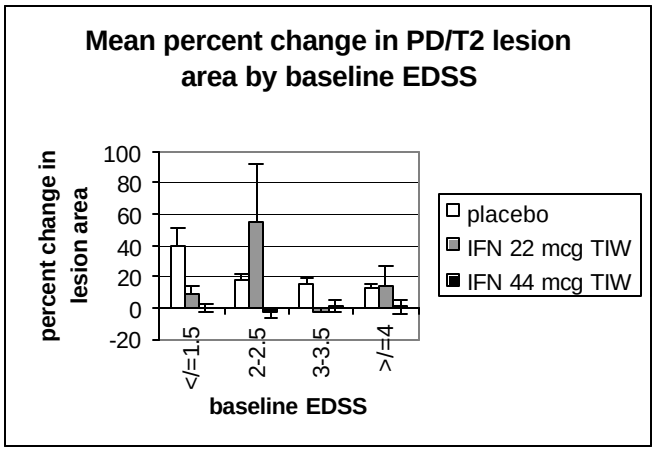
b) EDSS 3-month confirmed progressors



Numbers of subjects

baseline EDSS	placebo	IFN 22 mcg TIW	IFN 44 mcg TIW
≤1.5	61	54	56
2-2.5	51	49	47
3-3.5	43	45	46
≥4	25	34	30

c) Mean percent change in MRI PD/T2 lesion area



Numbers of subjects

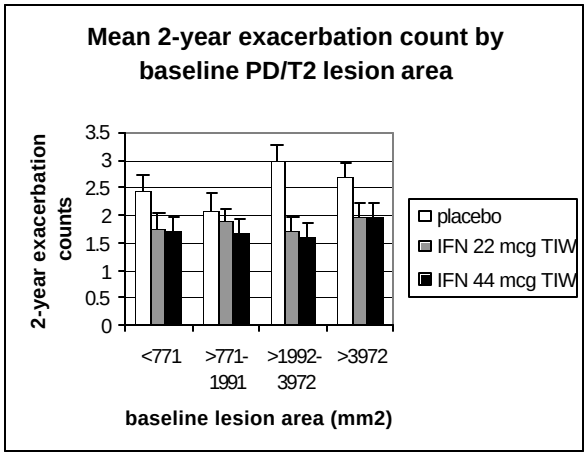
baseline EDSS	placebo	IFN 22 mcg TIW	IFN 44 mcg TIW
</=1.5	58	52	52
2-2.5	49	46	46
3-3.5	40	41	44
>/=4	25	32	29

Comment

The effects of both doses of IFN on the clinical measures were consistent with benefit at both extremes of EDSS. The effects of low-dose IFN on the MRI parameter was somewhat inconsistent, but high dose IFN provided consistent benefit.

Figure 10. Efficacy and imaging outcomes by baseline MRI PD/T2 lesion area (means shown with standard errors)

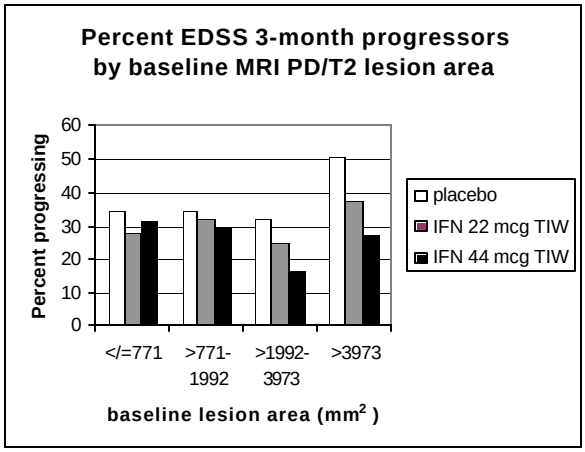
a) 2-year exacerbation counts



Numbers of subjects

baseline PD/T2 lesion area (mm^2)	placebo	IFN 22 mcg TIW	IFN 44 mcg TIW
≤771	48	51	42
>771-1991	40	45	54
>1992-3972	46	54	40
>3972	53	39	48

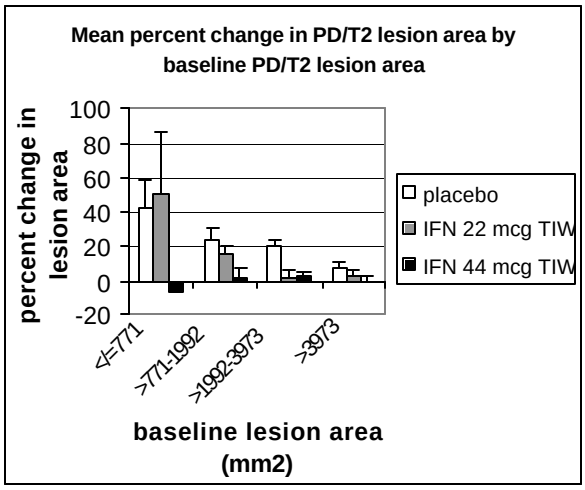
b) EDSS 3-month confirmed progressors



Numbers of subjects

baseline PD/T2 lesion area (mm^2)	placebo	IFN 22 mcg TIW	IFN 44 mcg TIW
≤771	47	51	41
>771-1992	38	44	54
>1992-3973	44	52	36
>3973	51	35	48

c) Mean percent change in MRI PD/T2 lesion area



Numbers of subjects

baseline PD/T2 lesion area (mm^2)	placebo	IFN 22 mcg TIW	IFN 44 mcg TIW
≤771	46	49	38
>771-1992	36	42	52
>1992-3973	44	49	35
>3973	46	31	46

Comment

There was clinical benefit from both doses of IFN at both extremes of MRI lesion area. The effects of low-dose IFN seemed to wane in the group with the most tendency toward MRI progression, those with the least MRI baseline PD/T2 lesion area.

Summary comments on exploratory analyses for differential benefit

These exploratory analyses failed to suggest critical thresholds of age, gender, weight, baseline EDSS, or baseline MRI PD/T2 lesion area outside of which the administration of IFN would reliably be expected not to confer benefit.

Antibody development and effect on efficacy

Table 18 summarizes the time course of development of neutralizing antibodies, and the numbers of subjects with a titer at the end of the trial. In all, 106 subjects developed some neutralizing antibody titer over the course of the trial (including one placebo subject who developed a titer at the end). Most antibody development occurred by the 18-month visit. Remarkably, there was more antibody development in the low-dose group compared to that in the high-dose group.

Table 18. Neutralizing antibody; time course of first development and end-of-trial antibody status (numbers of subjects)

visit	IFN 22 mcg TIW *n=189*	IFN 44 mcg TIW *n=184*
6 months	*23*	*26*
12 months	*16*	*13*
18 months	*17*	*6*
24 months	*3*	*0*
total	*59*	*45*
(+) at 24 months	*47*	*23*

Note:1 placebo subject developed neutralizing antibodies at 24 months;1 low-dose subject (+) at baseline

Table 19 shows exacerbation rates for those who developed any neutralizing antibody titer at the given time points, as presented by Serono.

Table 19. Exacerbation count by time of formation of any level of neutralizing antibody titer

	Placebo	IFN 22 mcg TIW				
Interval for count (months)	*n=187**	never titer *n=129*	(+) at 6 months *n=23*	(+) at 12 months n=*17*	(+) at 18 months *n=17*	(+) at 24 months *n=3*
0-6	0.81	0.60	0.35	0.29	0.41	1.00
6-12	0.67	0.42	0.57	0.47	0.53	1.00
12-18	0.55	0.37	0.39	0.59	0.59	0.67
18-24	0.55	0.41	0.39	0.29	0.53	0.67
0-24	2.56	1.80	1.70	1.65	2.06	3.33
	Placebo	IFN 44 mcg TIW				
Interval for count (months)	*n=187**	never titer *n=139*	(+) at 6 months *n=26*	(+) at 12 months *n=13*	(+) at 18 months *n=6*	(+) at 24 months
0-6	0.81	0.54	0.35	0.46	0.33	-
6-12	0.67	0.43	0.42	0.38	0.50	-
12-18	0.55	0.36	0.46	0.23	0.17	-
18-24	0.55	0.41	0.42	0.69	0.67	-
0-24	2.56	1.75	1.65	1.77	1.67	-

*Includes one subject who developed a neutralizing antibody titer

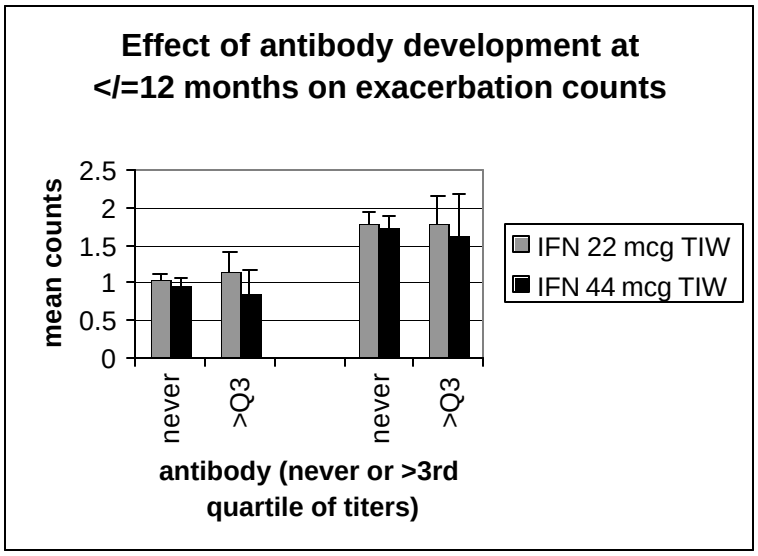
There is no significant effect of the development of antibody on subsequent efficacy. Serono's results for subjects developing high titer neutralizing antibody (using Serono's definition, i.e., ≥20 neutralizing units/ml) paralleled those for the development of any titer.

CBER sensitivity analyses of effect of antibody development on efficacy

1) Effect on exacerbation counts

To examine the effect of neutralizing antibody development further, CBER calculated total exacerbation counts at 1 and 2 years for those who never developed any neutralizing antibody titer and for those who developed a titer (at ≤ 12 months) greater than the 3rd quartile of titers for the low-or high-dose treatment groups. Figure 11 suggests that there is no differential efficacy after the development of high titer antibody.

Figure 11. Effect of the development of high titer antibody on exacerbation counts at 1 year (left) or the total at 2 years (right).



Numbers of subjects

Antibody status	IFN 22 mcg TIW	IFN 44 mcg TIW
never titer	129	139
>Q3	14	13

2) Effect on deterioration in disability

The percents of each group progressing in EDSS as a function of <u>any</u> antibody titer at any time was also examined. As Table 20 shows, there was no noticeable effect of antibody development, when all titers were considered as a group.

Table 20. Effect of the development of antibody titer on EDSS progression: percent (*n*) of treatment group

	Placebo (*n*)	IFN 22 mcg TIW (*n*)	IFN 44 mcg TIW (*n*)
never (+)	39% (*69/179*)	30% (*37/122*)	26% (*35/134*)
ever (+)	- (*1*)	28% (*17/60*)	29% (*13/45*)

3) Effect on MRI PD/T2 lesion area

Although there was no apparent loss efficacy in the exacerbation and disability parameters examined, an analysis of MRI PD/T2 lesions suggests that there may be a decrement in the effect of IFN on brain pathology in the presence of neutralizing antibody. This analysis is shown in Table 21.

Table 21. Effect of development of antibody on percent change in MRI lesion area

statistic		placebo	IFN 22 mcg TIW	IFN 44 mcg TIW
median				
	never (+)	10.9 *n=186*	-5.9 *n=129*	-6.3 *n=139*
	ever (+)	- *n=1*	9.1 *n=60*	8.6 *n=45*
mean (± sem)				
	never (+)	23.7 ± 4.6 *n=186*	16.3 ± 16.1 *n=129*	-5.0 ± 1.6 *n=139*
	ever (+)	- *n=1*	24.2 ± 5.4 *n=60*	12.6± 4.4 *n=45*

Comments

Exploratory analyses suggest that over the 2 years of the trial, clinical efficacy is preserved. However, the relative loss of efficacy in a laboratory measure of disease extent, PD/T2 lesion area, suggests that the development of antibodies was associated with a subclinical decrement in benefit. The clinical effects of antibody over a longer time period than 2 years may be greater than those seen over the 2 years studied in ----------.

Post-hoc analyses by Serono

Analyses of the following, not prospectively defined in the protocol, were presented by Serono:

1. Integrated disability score, defined as the area under the curve of a time-EDSS plot with respect to baseline. This measure, which is not in wide currency in the field of MS, provides a number that is difficult to assess, as changes in EDSS do not have uniform meaning throughout the scale.
2. Time to confirmed worsening in the 10-point ambulation index, defined as a worsening in the index by 2 steps confirmed 3 months later. Serono's analysis, a Cox proportional hazards analysis taking center into account, showed that only the high-dose group approached significance.
3. Change in an index of arm function (ref. 2). No statistically significant effect of treatment was seen.
4. Scripps neurological rating scale, change from baseline to end of study. The high-dose group approached significance in a difference from placebo, but the low dose group did not.

Comment

These measures are not used by the neurological community as benchmarks for the clinical course or effectiveness of therapy in MS, and should not at this time be used to alter the overall impression of efficacy.

Subgroup analyses performed by Serono (EDSS ≥4)

Serono presents an argument for the acceptance of an additional clinical category of MS: "transitional MS." According to Serono, this category disease occurs in patients at the end of a relapsing-remitting course and prior to a secondary progressive course. The clinical syndrome is described in the final report to -------- as "a new and aggressive stage of disease where [patients] begin to accumulate sequelae following their exacerbations," where there is "a marked increase in disease activity, baseline EDSS>3.5 (ranging from 4.0 to 8.0), and resistance to standard therapies.."

Other characteristics of "transitional MS" as mentioned in the preceding paragraph were not specific criteria for inclusion into this subset, although baseline characteristics such as MRI lesion area on PD/T2 scans, as well as the ambulation index and Scripps NRS, and measures of on-study disease activity such as mean numbers of exacerbations and hospitalizations, percentages of subjects with a confirmed change in EDSS, decreases in median time to first exacerbation, first quartile of time to first progression, and mean new PD/T2 lesions were higher than in the cohort with EDSS <4.

The following analyses were performed in the subgroup of subjects in this trial with EDSS ≥4 (94 subjects, approximately 1/6 of the entire subject pool).

Table 22. Efficacy endpoints for subjects with baseline EDSS ≥4.0 (*cohort size: 94*)

Endpoint	placebo	IFN 22 mcg TIW	IFN 44 mcg TIW	p-value IFN 22 vs placebo	p-value IFN 44 vs placebo	p-value IFN 22 vs IFN 44
mean exacerbation count at 2 years	3.07	1.83	1.22	0.02	0.0002	0.12
percent exacerbation-free subjects	7	18	32	0.39	0.05	0.18
time to first exacerbation (Q1, months; median, months)	1.2;2.8	1.8;7.5	2.9;10.8	0.06	0.005	0.25
mean number of moderate-severe exacerbations	1.8	1.0	0.9	0.09	0.05	0.67
percent of subjects with confirmed EDSS progression	56	41	27	statistics not presented		
time to confirmed EDSS progression (Q1, months; median, months)	7.3;21	7.5;not reached	21.3; not reached	0.45	0.05	0.22
mean hospitalizations for MS	0.86	0.54	0.48	0.12	0.15	0.94
number (%) subjects hospitalized	*13/26* (50)	*12/34* (35)	*11/30* (37)	statistics not presented		
mean steroid courses for MS	2.5	1.2	0.9	0.12	0.01	0.28
Number (%) subjects requiring steroids	*22/27* (81)	*21/34* (62)	*17/30* (57)	statistics not presented		
percent change in PD/T2 lesion area (median, mean)	5.4; 12.2	-2.3; 13.6	-6.9; 0.7	0.16	0.02	0.31
PD/T2 active lesions (median; mean)	1.9; 2.6	0.9; 1.7	0.5; 0.9	0.07	0.0002	0.03
% PD/T2 active scans (median; mean)	75; 71	50; 52	25; 36	0.04	<0.0001	0.02

Comment

These post-hoc analyses suggest that efficacy was maintained for most parameters listed; the dose trend seen in the overall subject pool was maintained, as well.

Summary (efficacy)

- Rebif caused statistically significant decreases in exacerbation counts that were robust to sensitivity and subgroup analyses. Time to exacerbation was also effected, but duration of exacerbations was not changed by treatment with Rebif. Differences in counts of moderate and severe exacerbations paralleled those in counts overall. Differences between treatments in exacerbation parameters trended toward a dose relation, with no statistically significant differences between active treatments.
- Other clinical outcomes as defined in the protocol (steroid use in MS, hospitalization rates for MS) were supportive of efficacy.
- Rebif caused statistically significant increases in the important supportive endpoint, time to disability. This effect was robust to sensitivity and subgroup analyses.
- Rebif caused decreases in MRI measures of MS disease pathology and activity. The differences between treatments in PD/T2 lesion areas were statistically significant, in distinction to changes seen in clinical parameters. However, differences between treatments in activity measures as determined in subgroups were not statistically different.
- While the analysis of considerably more subjects than projected by sample size estimates(560 as opposed to around 300) could have the effect of increasing the statistical significance of marginal results, it would not change the magnitude of the clinical effect and the overall conclusions of the trial.
- Post-hoc analysis of EDSS ≥4 group showed dose-related treatment trends.

RESULTS: SAFETY

Deaths

There were two deaths. One placebo-group subject committed suicide and one low-dose subject died following a fall, with closed head injury, after 12 months of treatment. Neither event was considered related to the trial medication.

Adverse events: Serious adverse event reports

The numbers of serious adverse events reported differed among the clinical trial final report, the electronic data base, and the integrated summary of safety. For the following analysis, case summaries were examined and counted as individual serious adverse events. Certain events, i.e., the onset of pregnancy and illnesses or laboratory abnormalities treated on an outpatient basis, were excluded because they did not meet the regulatory definition of "serious". In some instances, more than one clinical event was recorded in a report (separated by the "/" symbol in Table 21). In rare reports readmissions occurred for the same event (not captured in this tabulation); there was no pattern of this reporting characteristic occurring among the treatment groups, nor did it occur for a certain clinical event predominantly.
Excluding hospitalizations for MS (reported in "Efficacy" above), serious adverse events occurred in 25, 25, and 19 subjects in the placebo, low-, and high-dose groups, respectively. Table 23 shows the organ systems affected in the serious adverse events reports, using the

electronic data base as the source for the WHOART organ system classification (it was not provided with the case summaries).

Table 23. Numbers of serious adverse event reports, organ system affected (WHOART definitions)

WHOART organ system	Placebo *n=187*	IFN 22 mcg TIW *n=189*	IFN 44 mcg TIW *n=184*
gastrointestinal	*5*	*3*	*0*
body as a whole/urinary system	*1*	*0*	*0*
heart rate and rhythm	*1*	*0*	*0*
liver and biliary system	*1*	*0*	*0*
not classified	*2*	*4*	*3*
psychiatric disorders	*3*	*5*	*2*
female reproductive	*5*	*1*	*1*
respiratory	*1*	*1*	*2*
respiratory/urinary system	*1*	*0*	*0*
secondary terms	*5*	*4*	*6*
secondary terms/psychiatric	*1*	*0*	*0*
skin and appendages	*2*	*1*	*0*
urinary system	*1*	*1*	*0*
body as a whole-general	*0*	*2*	*2*
central and peripheral nervous system	*0*	*2*	*0*
platelet, bleeding, and clotting	*0*	*2*	*1*
resistance mechanism	*0*	*1*	*0*
resistance mechanism/platelet, bleeding, and clotting/heart rate and rhythm	*0*	*1*	*0*
application site	*0*	*0*	*1*
neoplasm	*0*	*0*	*1*
vascular (extracardiac)	*0*	*0*	*1*
vision	*0*	*0*	*1*
white cell and RES	*0*	*0*	*1*
TOTAL	*29*	*28*	*22*

Below is a discussion of the events that occurred in the active groups only.

Serious adverse events pertaining to the gastrointestinal system occurred in 2 subjects in the low-dose group. One subject had an appendectomy and treatment of a renal calculus; another had a spastic colon.

The events listed as "not classified" had no WHOART term in the data base; these events were of diverse natures, and all occurred in different subjects. For the 1) low- and 2) high-dose group these events were: 1) resection of ovarian tumor, diagnostic laparotomy for suspected endometriosis, diagnostic tests for suspected neuroborreliosis, and obstipation; and 2) elective knee surgery, induced abortion, and removal of previously placed pin from limb. The case report from the induced abortion states "Examination of the fetus revealed a circulatory disturbance and a placental infarction." The meaning of this statement is unclear at this time.

"Psychiatric disorders" were depression and suicide-related. These were reported in 3, 4, and 2 subjects respectively in the 1) placebo, 2) low- and 3) high-dose groups. The events were:

1) suicidal ideation (3 occurrences); 2) acute depression, worsened depression, suicidal ideation (2 occurrences) and suicide attempt; 3) suicidal ideation and suicide attempt. These are also discussed below under "Psychiatric events, suicide, and suicide-related events." It should be noted that the "serious" events considered here do not include all the events examined in that discussion.

The female reproductive-related serious events in the active treatment groups were 1 diagnostic laparoscopy and hysteroscopy for suspected endometriosis and 1 hysterectomy and ovariectomy for an ovarian cyst (1 subject in each active group). Two additional subjects in the high-dose group had female reproductive-related events, classified as secondary terms: removal of an ovarian cyst with dilation and curretage, and a hysterectomy for bleeding.

Respiratory system disorders occurring in active treatment groups were 1 subject with bibasilar pneumonia after a Hemophilus influenzae infection in the low-dose group, and 2 subjects, one with an asthma exacerbation and one with a viral URTI, in the high-dose group. An additional subject in the high-dose group had a varicella pneumonia, classified as a secondary term (immediately below). It should be noted that infections as a whole (see discussion of adverse events of all severities) were not increased in either incidence or severity.

Secondary terms were of various natures, with no predominating type. For the 1) low-dose, and 2) high-dose subjects they were: 1) closed head injury following a fall (resulting in death), laminectomy (2 subjects), reconstruction of ankle following a fall; and 3) removal of ovarian cyst with dilation and curretage, hysterectomy for bleeding, reduction of fracture following a fall, removal of knee cartilage, enlargement of nasal sinuses, and varicella pneumonia.

Body as a whole disorders were of various natures, and occurred in 2 subjects in the 1) low- and 2 in the 2) high-dose groups. They were: 1) chest pain with anxiety and pyeloplasty and 2) elective nasal polypectomy and decompression of carpal tunnel constriction.

Two subjects, in the low-dose group, had serious nervous system events, of different natures: 1 had ataxia, leg weakness, and dysdiadokinesia, and the other had migraines.

Platelet, bleeding, and clotting disorders were found only in the active-treated groups. One low-dose subject had a venous embolus (pulmonary) and one subject in each active group had an arterial embolus (arm and hand). The two arterial cases had predisposing causes for the development of clots; one had been off IFN for over 16 months.

One subject, in the high dose group, had a serious application site event, an abscess that required surgical treatment. It should be noted that another subject, also in the high dose group, had serious infective lymphadenopathy (mentioned below).

Although there is one listing of "neoplasm" among the serious adverse events, a subject in the high-dose group with a resection of a Duke's B adenocarcinoma, there were other cancers treated throughout the groups in the trial, 2 in the placebo, and 1 in the low-dose group.

The white cell and RES-related serious adverse event, in a high-dose subject, was infective lymphadenopathy from injections (see discussion of application site event above). Additional events that occurred in single low- or high-dose subjects were:

<u>low-dose</u>

skin and appendages disorder: treatment of erysipelas

urinary system disorder: urinary bladder surgery to correct post-voiding residuals

resistance mechanism disorder: viral syndrome

"resistance mechanism/platelet, bleeding, and clotting/heart rate and rhythm"

event: urosepsis with septicemic shock, DIC, and cardiac arrest

high-dose

extracardiac vascular disorder: embolization of a newly MRI-diagnosed, asymptomatic cerebral aneurysm.

vision: cataract removal

Comments on serious adverse events

The small numbers of events of any one category makes strong associations with IFN difficult to assign. Nevertheless, it is worth noting that abscess and infective lymphadenopathy occurred only in treated groups, and are events in a spectrum of injection-site events that occurred with greater frequency in active-treated groups (see below). The pulmonary embolus and one of the arterial emboli were classified as life-threatening; thromboembolic events occurred only in active-treated subjects. Depression and related events did not occur with greater frequency among treated groups, but their possible association in some treated patients is concerning (see discussion of "Psychiatric" events below). The nature of the abnormality seen in the case of the induced abortion should be elucidated, but due to concerns over IFN treatment in pregnancy is worthy of mention.

Events leading to treatment discontinuation

Table 24 shows the reasons for treatment discontinuation when coded by Serono as an adverse event, "subject decision," or "other," using the following rules:

1) due to adverse events as tabulated by Serono

2) due to "subject decision" or "other," with a clear adverse event in the recorded comment (e.g., they do not include worsening of MS). These are denoted by an asterisk.

Table 24. Adverse experiences leading to treatment discontinuation

Treatment group	WHOART term (*"subject decision" or "other")
placebo	
	depression
	*"depression with suicidal thoughts"
	headache
	*"problematic side effects as judged by the patient"
	*"gynecological problems played a psychological role"
	*"occurrences of relapses and infections"
IFN 22 mcg TIW	
	depression
	depression
	*"severe psychological problems and reactive depressive mood"
	fever
	septicemia, DIC, cardiac arrest
	anaphylactoid reaction
	*"current medical condition-lymphopenia"
	injection site reaction
	*"painful injections"
IFN 44 mcg TIW	
	depression
	*"depressed and felt unmotivated to inject himself"
	*"very irritable, suicidal thoughts"
	suicide attempt
	injection site reaction
	fever, muscle pain, leukopenia, SGOT increased
	SGOT increased
	lymphopenia
	colon carcinoma
	bradypsychic responses
	palpitation
	*"intolerable side effects"

The number of discontinuations was increased in a dose-related manner. This trend was also evident in the number of dosing interruptions noted as protocol violations. The number of discontinuations due to depression and depression-related symptoms was mildly increased in the high-dose group, but due to the small number of events a clear dose relation is hard to establish. In general there were increased discontinuations from the active treated groups, but the numbers of discontinuations due to any one cause was small.

In addition to discontinuations due to adverse events, subjects discontinued for other reasons: the onset of pregnancy was the cause for treatment discontinuation in 1, 2, and 3 of the placebo, low-, and high-dose subjects; progression of disease in 3, 1, and 0; and protocol violation in 0, 0, and 1 subject. Of the 6 subjects who discontinued due to pregnancy, 5 had

healthy children, and one, a subject in the high-dose group, had an induced abortion (discussed above in serious adverse events).

Adverse events of all degrees of severity

Table 25 shows events that occurred with ≥2% increase in frequency in the high-dose group compared to placebo. Events with a ≥2% increase in the low-dose but not the high-dose groups showed small increases in frequency over placebo and were sporadic in nature.

Table 25. Percents of each group with adverse events that showed a ≥2% increase in the high-dose group compared to placebo

WHOART preferred term	Placebo *n=187*	IFN 22 mcg TIW *n=189*	IFN 44 mcg TIW *n=184*
headache	62.6	64.6	70.1
influenza symptoms	51.3	56.1	58.7

fatigue	35.8	32.8	41.3
fever	15.5	24.9	27.7
rigors	5.3	6.3	13.0
chest pain	5.3	5.8	7.6
malaise	1.1	4.2	5.4
injection site inflammation	15.0	65.6	65.8
injection site reaction	13.4	31.2	34.8
injection site pain	14.4	20.1	22.8
injection site mass	0.5	3.2	3.8
injection site necrosis	0.0	1.1	3.3
injection site abscess	0.0	1.1	2.2
abdominal pain	17.1	22.2	19.6
mouth dry	1.1	0.5	4.9
myalgia	19.8	24.9	25.0
back pain	19.8	23.3	24.5
insomnia	21.4	19.6	23.4
somnolence	0.5	3.7	4.9
rash maculopapular	1.6	4.8	4.3
lymphopenia	11.2	20.1	28.8
leukopenia	3.7	12.7	22.3
lymphadenopathy	8.0	11.1	12.0
granulocytopenia	3.7	11.6	15.2
wbc abnormal numbers	0.5	3.2	2.7
micturition frequency	3.7	1.6	6.5
vision abnormal	7.0	7.4	13.0
SGPT increased	4.3	19.6	27.2
SGOT increased	3.7	10.1	17.4
hepatic function abnormal	1.6	3.7	9.2
thrombocytopenia	1.6	1.6	8.2
thyroid disorder	3.2	4.2	6.0
anemia	2.7	2.6	4.9

Specific adverse events

Most of the events fell into the following categories: flu-like symptoms, injection site reactions, hematopoietic abnormalities, and elevations of liver enzymes. They do not represent new toxicities of IFN. These events will be reviewed separately below. In addition, psychiatric (and depression-related) and other, selected adverse events will be discussed.

Influenza-related events

The 2-year duration of the trial made it possible that influenza-like events could have originated from illnesses as well as the effects of IFN. The relative increase in frequency in these events may have been blunted. Increases for most influenza symptoms were evenly distributed among mild, moderate, and severe events for most influenza-related symptoms**,** with small increases in moderate reactions among treated subjects were seen in the categories "influenza symptoms" and "malaise" (rates in the placebo, low- and high-dose groups of 19.8,

24.3, 23.9 and 0, 2.6, and 3.8, respectively). Severe influenza-related symptoms in all groups were rare.

Injection site reactions

Injection site reactions are a clinical concern with the subcutaneous use of IFN; they occurred more frequently in the active-treated patients in the trial examined for the licensure of another IFN-β, Betaseron. In the current trial, most of the increase in injection site inflammation occurred in cases of mild severity (mild cases in placebo, low-, and high-dose groups occurred in 14.4, 61.4, and 60.9 percent of subjects, moderate in 1.1, 6.9, and 11.4 percent of subjects; severe in 0, 0.5, and 2.2 percent of subjects). However, as can be seen in the table above, necrosis and abscess occurred only in active-treated subjects (one case of abscess and one case of infective lymphadenopathy required hospitalization); in addition, one case of atrophy was reported, in a low-dose subject. Injection site bleeding occurred with greater frequency in active treatment groups: 0.5%, 2.1%, 2.2% of the placebo, low, and high-dose groups respectively. (In contrast, the incidence of injection site bruising was lower in the active groups; 11.2%, 3.7%, and 5.4% of the placebo, low, and high-dose groups, respectively.) All 8 cases of necrosis resolved and patients continued with study treatment. Three patients, 2 from the low-dose and 1 from the high-dose group, permanently stopped treatment due to injection site events.

Serono calculated the durations of "injection site reaction (other)" among all the subjects, assuming that an ongoing event would be captured, as it should have been, in consecutive 3-month visits. "Injection site reaction (other)" refers to events that did not fit into the established categories or that had different terms in the same report. Events were categorized as <1 month, 1-2 months, 3,6,9,12, or >12 months. Overall, the number of subjects with reactions of all durations increased with treatment:18 placebo, 60 low-dose, and 64 high-dose subjects reported this adverse event. The majority of subjects in treated groups had reactions of 3 months' duration (placebo, 6 cases, low-dose 30 cases, and high-dose, 36 cases). According to this analysis by Serono, there was an increase in very long cases among the active-treated subjects as well: 1 placebo, 2 low-dose, and 5 high-dose subjects experienced "injection site reaction (other)" for greater than 12 months. In addition, Serono reports that 2 subjects in the low-dose group and 5 in the high-dose group reported injection site reactions "off-and-on for more or less the entire study duration."

In sum, although there may be risks attendant to any subcutaneous injection, there were more significant reactions among the IFN-treated groups.

Hematopoietic abnormalities

Low- and high-dose IFN treatment caused an increase in subjects reporting lymphopenia. The increase tended to be more in mild cases in the low-dose subjects; however, the increase in lymphopenia in the high dose group was also associated with an increase in the severity of cases (rates for mild lymphopenia in placebo, low-, and high-dose groups were 7, 11.1, and 10.3; moderate lymphopenia, 4.3, 6.3, and 16.8; and severe lymphopenia, 3.7, 4.8, and 11.4). In addition, 2 subjects, in the high-dose group only, had asymptomatic high-grade lymphopenias that were classified as life-threatening. Although there were no cases of symptoms associated with lymphopenias, the increase in number of lymphopenias in both active groups, and in severity of lymphopenias in the high dose group, is cause for caution in the use of IFN. Increases in granulocytopenia were occurred primarily in mild and moderate cases (rates for mild granulocytopenia in placebo, low-, and high-dose groups were 3.2, 8.5, and 12.0; moderate

granulocytopenia, 0, 3.2, and 9.8; and severe granulocytopenia, 0.5, 0, and 1.1). The increase in thrombocytopenia and anemia, in the high dose group only, occurred as mild cases. These events are also reason to use IFN with caution.

Liver enzyme elevations

SGOT and SGPT increases paralleled each other and were primarily in cases of mild severity ; severe SGOT and SGPT abnormalities were rare. For SGOT rates of mild events in placebo, low-, and high-dose groups were 3.7, 6.3, and 15.2; moderate events, 0.5, 3.7, and 3.8; and severe events, 0.5, 1.6, and 0.5. For SGPT rates for mild events in placebo, low- and high-dose groups were 3.7, 13.2, and 23.4; moderate events, 1.1, 6.9, and 5.4; and severe events, 0.5, 1.6, and 1.1. There were very few hepatic-related clinical events in the data base, and very few clinical comments attached to these laboratory elevations.

Selected miscellaneous events

The incidence of infections, including viral infections such as herpes zoster and herpes simplex, was not increased among active-treated subjects. Severity of infections overall was not increased. However, the increased incidence of cytopenias is a reason for extra vigilance for infections in the chronic administration of IFN.

Increases in the event "thyroid disorder" occurred in cases of mild and moderate severity (rates for mild thyroid disorder in placebo, low-, and high-dose groups were 2.7, 2.6, and 4.9; moderate thyroid disorder, 0.5, 1.6, and 1.1). Active treatment was associated with both increases and decreases in TSH in small numbers of subjects (the numbers of subjects whose values increased from normal to high in placebo, low-, and high-dose groups were 0, 2, and 5, and the numbers of subjects whose values decreased from normal to low were 3, 2, and 6). In the subjects affected by a thyroid disorder, active treatment was associated with a small increase in incidence of subjects with increases in thyroid stimulating hormone. One placebo subject had mild hypothyroidism and 1 subject in the high-dose group had severe hypothyroidism. The etiology of thyroid disorders in this trial is unclear, but there is a suggestion that thyroid monitoring is indicated with IFN treatment.

High-dose treatment was associated with the adverse event coded by the WHOART preferred term "vision abnormal." These events were of mild to moderate severity. None of the subcategories of "vision abnormal" listed (eye pain, conjunctivitis, diplopia, or xerophthalmia) showed the same trend toward an increase in incidence in the high dose group as the parent category . It is of interest that in this blinded trial, of the 67 incidences of "vision abnormal" as a WHOART preferred term, 7 were considered possibly or probably related to IFN treatment, 6 of which occurred in active-treated subjects. Since MS activity can be elicited by increased temperature, an association with fever was sought: it was not possible, from the data provided, to attribute blurred vision to the adverse event "fever" in the majority of cases. Serono reports a single case of a healthy subject who received Rebif in a previous study, with the subsequent onset of blurred vision, which was presumptively diagnosed as optic neuritis. Thus, although blurred vision is a symptom of MS, the higher incidence of abnormal vision in the high dose group warrants vigilance.

The adverse event "seizure" was recorded in the electronic data base for 2 subjects, both in the low-dose group. Both subjects had a history of seizures, and one of the cases was associated with starting a course of fluoxetine. Serono states that their clinical review of the data from trial --------- revealed seizures in1 subject in the placebo group, 3 in the low-dose group,

and none in the high-dose group. The placebo subject had focal seizures, and the 3 low-dose subjects had generalized seizures. Because of the inconsistency between Serono's account of the numbers of seizures and the electronic data base, further analysis of these events is warranted. The evidence in this trial does not eliminate the concern over the relationship between IFN treatment and seizures.

One subject in the low-dose group experienced a severe anaphylactoid reaction that was deemed probably due to the administration of IFN. This reaction occurred after approximately 1 month of treatment; it progressed over several days to "localised symptoms of uriticaria, generalised pruritis, and swollen red scars at the injection site." The subject recovered after treatment. According to the adverse reaction report, subsequent skin tests showed that the subject was allergic to a component in both the active drug and placebo. In light of a recent report of anaphylaxis in a patient receiving rIFN-β1a ("Severe anaphylaxis with recombinant interferon beta" in Neurology (January, 1999) more information on the incidence of anaphylaxis with the use of Rebif needs to be collected and analyzed.

Psychiatric events, suicide, and suicide-related events

Because of the concern with the association of MS and the administration of IFN with the occurrence of depression and suicide, psychiatrically-related events and suicide and its related events were examined in more detail. Table 26 shows the numbers of subjects experiencing psychiatrically coded events.

Table 26. WHOART psychiatric disorders, percent (number)of subjects

WHOART preferred term	Placebo *n=187*	IFN 22 mcg TIW *n=189*	IFN 44 mcg TIW *n=184*
depression	27.8 (*52*)	20.6 (*39*)	23.9 (*44*)
moderate depression	10.2 (*19*)	6.9 (*13*)	10.9 (*20*)
severe depression	2.1 (*4*)	2.1 (*4*)	1.1 (*2*)
lifethreatening depression	1.1 (*2*)	0.5 (*1*)	0.5 (*1*)
insomnia	21.4 (*40*)	19.6 (*37*)	23.4 (*43*)
anxiety	5.9 (*11*)	4.8 (*9*)	7.6 (*14*)
nervousness	6.4 (*12*)	5.3 (*10*)	6 (*11*)
anorexia	3.7 (*7*)	4.8 (*9*)	3.3 (*6*)
somnolence	0.5 (*1*)	3.7 (*7*)	4.9 (*9*)
sleep disorder	2.1 (*4*)	3.7 (*7*)	2.2 (*4*)
emotional lability	3.2 (*6*)	2.1 (*4*)	1.1 (*2*)
amnesia	1.6 (*3*)	2.1 (*4*)	1.1 (*2*)
agitation	0.5 (*1*)	0.5 (*1*)	2.2 (*4*)

The psychiatric events listed above were not increased among the active treated subjects, with the exception of somnolence. Table 25 shows adverse events related to suicide. Adverse event descriptions were reviewed, as the WHOART preferred terms for these events partially overlap. For the Table 27 (below), suicidal ideation includes all events having to do with suicide, and the suicide was not counted as an attempt.

Table 27. Numbers of subjects with suicidal ideation, suicide attempt, and suicide.

Adverse event description	placebo	IFN 22 mcg TIW	IFN 44 mcg TIW
suicidal ideation	3	3	4
suicide attempt	1	1	1
suicide	1	0	0

There was no increase among active-treated subjects in the incidence of suicide or suicide-related events.

Depression-related serious adverse event reports

There does not seem to be a relationship of treatment with the overall incidence of depression or suicide. However, perusal of serious adverse event reports suggests that concern is warranted. One case report in an active-treated subject was particularly concerning: it states that on the day of medication initiation, "..the patient started study medication and from this date reported a mood change/depression, which worsened… the patient was withdrawn from the study due to depression and suicidal ideation, after one month of study medication. The depressive episode resolved completely 2 weeks after drug cessation.." This event was deemed "possibly" related to study medication. Although none of the serious depression reports among active-treated subjects had a causality attribution of greater than "possible," and there were "possible" attributions in the placebo as well as each treatment group, the association of IFN treatment with depression cannot be ruled out.

Pyschological testing for depression

The trial included 3 self-administered psychological tests for depression, conducted biannually, in English-speaking countries only. These were the Beck's Hopelessness Scale (BHS), the Centre for Epidemiologic Studies Depression Mood Scale (CES-D), and the General Health Questionnaire (GHQ). The BHS was developed to measure symptoms of hopelessness or pessimism; the CES-D, to distinguish depression (some of the items derive from the BHS, so it is not an entirely independent measurement tool), and the GHQ, to detect the inability to carry out normal functions and the appearance of new, distressing psychological events, not necessarily depressive ones.

Table 28 shows numbers of subjects with scores greater than 8 on the BHS, i.e., moderate or higher degrees of hopelessness. In this test, scores from 0-3 represent normal to asymptomatic, and 4-8, 9-14, and >14 represent mild, moderate, and severe degrees of hopelessness, respectively. Numbers of subjects with scores >14 were too small for comparisons of dose- and time-related trends.

Data from the BHS derived from 262 subjects (this included 1 Swiss subject). The number of subjects enrolled at the sites from which the data are reported, which excluded Sweden, Netherlands, Belgium, Finland, Denmark, and Switzerland, was 266.

Table 28. Moderate or higher degrees of hopelessness, percents of subjects taking the Beck's Hopelessness Scale questionnaire. The numbers of subjects taking the questionnaire are in parentheses.

	Placebo	IFN 22 mcg TIW	IFN 44 mcg TIW
baseline	6% (*80*)	14% (*77*)	14% (*79*)
6 months	16% (*77*)	8% (*73*)	11% (*72*)
12 months	13% (*69*)	8% (*65*)	10% (*69*)
18 months	13% (*71*)	12% (*74*)	9% (*75*)
24 months	9% (*78*)	13% (*76*)	9% (*78*)

Overall results were similar for the other two tests, i.e., baseline scores were at or near the normal ranges, with no dose-related trend toward an increase numbers of subjects with depressive scores among treated subjects.

Comments

These data on psychological testing, in conjunction with the data on serious depression and its related events, suggest that serious depression-related sequelae of IFN administration may occur rarely, in susceptible individuals with MS, but that overall the risk of depression and depression-related events is not increased over that of placebo-treated subjects.

Unscheduled visits

Numbers of unscheduled visits were increased in the placebo group: in the placebo, low-, and high-dose groups the mean number of unscheduled visits per subject was 1.44, 1.04, and 0.88, respectively. In addition, there was a larger number of subjects with no unscheduled visits on active treatment: 37%, 56%, and 54% in the placebo, low-, and high-dose groups, respectively. While the outcome of unscheduled visits (i.e., whether they resulted in treatment for MS) was not recorded, these data suggest that there were more clinically significant medical problems among the placebo than among the treated groups.

Concomitant medication use

Serono's tabulation of concomitant medications provided data on use at baseline and during treatment, and thus doesn't elucidate any differences between the baseline and treatment periods. Serono provided, upon request, a tabulation of the numbers of subjects taking selected medications at least once, comparing the baseline and the treatment periods. These medications included the major symptomatic treatments for spasticity, paroxysmal symptoms (e.g., trigeminal neuralgia, tonic spasms), fatigue, urinary urgency or hesitancy, as well as antidepressants. Oral and parenteral steroid use for all indications, not only for MS exacerbations, was included as well. Baseline usage of all groups of medications was similar among the three treatment groups (except in the case of medications grouped for their use in paroxysmal symptoms, in which there were 8, 12, and 15 subjects in the placebo, low, and high-dose groups, respectively).

Comment

The numbers of subjects who used at least one MS-related medication in each group increased during the trial, but the increase was either equal among the groups or greater for placebo. Thus there was no indication that a deleterious effect of active treatment was masked by the use of concomitant medications for MS. However, at the time of the writing of this review, Serono has not provided a tabulation of all medications separated by baseline use and during-trial use. Thus this part of the review is incomplete.

Summary (safety)

- The incidence of mild-to-severe depression, suicide, and suicidal ideation was not increased by active treatment. However, the evidence from the trial is not enough to rule out the possible association of IFN treatment with serious depression.
- Hematologic and hepatic toxicity was noticeably increased in a dose-related manner. Some of these abnormalities resulted in decisions to stop treatment. Although severe hematopoietic events occurred with greater frequency in active treated groups, there was no increase in infections. Hepatic enzyme abnormalities appeared to be isolated laboratory abnormalities.
- There were noticeable increases of important injection site adverse events with Rebif therapy. Rare discontinuations of treatment occurred related to these events.
- There was an increase in active-treated subjects in the known flu-like symptoms of interferon treatment.
- Mild to moderate diverse thyroid abnormalities and abnormal vision occurred in high dose subjects, the etiologies of which were not clear from the data provided.
- Concomitant medication for the symptoms of MS did not suggest a masking of adverse effects of Rebif and did not raise safety concerns. However, at the time of the writing of this review, there are no data on the use of other concomitant medications separated by baseline and intertrial periods, so this part of the review is incomplete at this time.

TRIAL ----------

Trial ---------- is the only other completed trial in patients with RRMS. Because of it was small and of open-label design, conclusions drawn from it are tentative, so it will be reviewed only briefly here.

DESIGN AND ANALYTICAL PLAN

The protocol for this trial was not received for comment by CBER. Seventy-two subjects were recruited into 2 centers in Italy in this open-label trial from June 1993 to January 1994. They were aged 15-45, with a mean disease duration of 5 years, and a mean EDSS of 2.1. After an observation period of 6 months, 68 subjects (there were 4 dropouts during observation) were randomized to treatment with reconstituted, lyophilized IFN either 11 or 33 mcg SC three times weekly for an additional 6 months. The primary endpoints, number and volume of enhancing lesions on MRI, was based on monthly MRI performed with and without the administration of gadolinium. Neuroradiologists were blinded to the dates of the exams upon analysis. Secondary endpoints were numbers of new and enlarging lesions on MRI and the number of clinical relapses.

RESULTS: EFFICACY

Table 29 shows Gd-enhancing lesion counts and volumes.

Table 29. MRI lesions measures per month per subject (mean ± sd)

		IFN 11 mcg TIW	IFN 33 mcg TIW	all subjects
Total lesion count (1^o endpoint)	observation period	3.47 ± 5.02 *n=35*	2.42 ± 3.45 *n=33*	2.96 ± 4.33 *n=68*
	treatment period	1.77 ± 2.62 *n=35*	0.86 ± 1.68 *n=33*	1.33 ± 2.24 *n=68*
	p-value*	<0.001	<0.001	<0.001
Volume of Gd-enhanced lesions (mm^3)(1^o endpoint)	observation period	558 ± 821 *n=35*	379 ± 522 *n=33*	471 ± 693 *n=68*
	treatment period	220 ± 357 *n=35*	101 ± 184 *n=33*	162 ± 290 *n=68*
	p-value*	<0.001	<0.001	<0.001
T2 new lesions (2^o endpoint)	observation period	5.68 ± 5.56 *n=34*	3.94 ± 5.01 *n=33*	4.82 ± 5.33 *n=67*
	treatment period	1.97 ± 2.59 *n=34*	1.18 ± 1.88 *n=33*	1.58 ± 2.28 *n=67*
	p-value*	<0.001	<0.001	<0.001
T2 enlarging lesions (2^o endpoint)	observation period	2.27 ± 2.35 *n=34*	1.82 ± 2.33 *n=33*	2.05 ± 2.33 *n=67*
	treatment period	0.97 ± 2.10 *n=34*	0.46 ± 0.87 *n=33*	0.72 ± 1.62 n=*67*
	p-value*	0.001	0.004	<0.001

*Wilcoxon signed rank test for within-dose comparison of treatment period to observation period; Student's t-test for all-subject comparison

Table 30 shows exacerbation rates (a secondary endpoint).

Table 30. Exacerbation rates per period of the trial (mean ± sd)

	IFN 11 mcg TIW *n=35*	IFN 33 mcg TIW *n=33*	all subjects *n=68*
observation period	0.91 ± 0.98	0.79 ± 0.93	0.85 ± 0.95
treatment period	0.43 ± 0.70	0.24 ± 0.50	0.34 ± 0.61
p-value*	0.007	0.003	<0.001

*Wilcoxon signed rank test for within-dose comparison of observation period to treatment period; Student's t-test for all-subject comparison

Summary (efficacy)

These results, suggestive only as they come from an open-label, uncontrolled trial, support the efficacy conclusions of the controlled trial. Mean volume of Gd-enhancing lesions and other, nonprotocol-defined parameters submitted by Serono support the primary endpoint: numbers of prednisolone pulses to treat exacerbations, numbers of exacerbation-free subjects, and measures of GD-enhancing lesion number and volume related to corticosteroid use. Mean values of EDSS and the ambulation index score did not change in a clinically meaningful or statistically significant way for any group.

RESULTS: SAFETY

Serono reports data on 68 subjects, i.e., those who did not drop out during the observation period. In general very few adverse experiences occurred during the observation period.

There were no deaths. Two serious adverse events were reported: 1 subject (11 mcg TIW) dropped out of treatment due to cognitive deficits diagnosed as an MS exacerbation, an event that resulted in hospitalization; 1 subject was hospitalized for treatment of pneumonia. Both subjects improved. One subject had a dose reduction (33 to 11 mcg TIW) due to asthenia. A temporary discontinuation of treatment occurred once in each dose group (dyspnea as a cause in the low dose group, bronchopneumonia in the high dose group). Table 31 shows the numbers of subjects with various adverse experiences if experienced by ≥2 subjects in either treatment group.

Table 31. Trial GF6613: Percents (*n*) with adverse events, among adverse events that occurred in ≥2 subjects per treatment group during the treatment period.

	observation period	11 mcg TIW *n=35*	33 mcg TIW *n=33*
flu-like symptoms	0	31 (*11*)	42 (*14*)
headache	1 (*1*)	23 (*8*)	30 (*10*)
injection site reaction	0	40 (*14*)	61 (*20*)
fever	1 (*1*)	17 (*6*)	24 (*8*)
asthenia	0	14 (*5*)	15 (*5*)
nausea	0	0	9 (*3*)
pain	4 (*3*)	17 (*6*)	13 (*9*)
insomnia	1 (*1*)	9 (*3*)	12 (*4*)
depression	1 (*1*)	6 (*2*)	6 (*2*)
irritability	0	3 (*1*)	9 (*3*)
myalgias	0	6 (*2*)	6 (*2*)
URTI	15 (*10*)	0	9 (*3*)
menstrual disorders*	0	8 (*2*)	5 (*1*)
cystitis	3 (*2*)	3 (*1*)	6 (*2*)

*among premenopausal females

No severe adverse experiences were reported. There were no suicides or suicide attempts. Grade I leukopenia occurred in 9% of each group, but was already present in 2 of the lower dose group and 1 in the higher dose group. Increases in AST were seen in 12% and 9% of the high and low dose groups; increases in AST were seen in 14% and 26% of subjects per group.

Summary (safety)

The types of adverse experiences and their frequencies were not substantially different from those seen in the controlled trial; these data do not change the understanding of Rebif gained from that trial.

POST-MARKETING ADVERSE EVENT REPORTING

Rebif was marketed only in Italy at the time of the compilation of the post-marketing report. Serono reports that a single event, leg thrombophlebitis, has been reported post-marketing up to the cutoff date of the Integrated Summary of Safety (December 31, 1996). The investigator reported the event as "unlikely" due to Rebif, and the patient recovered.

SAFETY EXPERIENCE USING REBIF IN OTHER MS TRIALS

Serono submitted an overall summary of safety in various MS disease forms as of December 31, 1996. The numbers of subjects treated with Rebif in various indications is estimated in some cases due to the blinded nature of ongoing investigations: in relapsing-remitting MS, an estimated total of 500 subjects (not including ---------), in secondary progressive MS, a total of 565 subjects, and in other demyelinating disorders ("clinically probable or laboratory-supported" MS and chronic inflammatory demyelinating polyneuropathy), a total of 126 subjects. The gender ratio for the trials (omitting several very small trials) was similar to that of the general MS population (2:1, female:male); the great majority (99%) of subjects in 2 trials comprising 911 subjects were Caucasian (information is not

shown for the majority of trials). The age range (trials including ----------) was 12-65, with 95% of the population between 20-55 years of age.

This review will focus on safety in the MS population in trials other than ---------- and ---------- (reviewed above). The number of subjects in RRMS indications is estimated at 510 (632 subject-years of observation); in secondary progressive indications 894 (1402 subject-years of observation); in other MS forms and demyelinating diseases, 249 (136 subject-years of observation). The data provided include deaths, dropouts due to adverse events, and other serious events.

There were 2 deaths in RRMS trials other than ------- and 4 other deaths, all in SPMS. Two of the deaths were due to suicide (both in SPMS, blinded treatment assignment); the other deaths in RRMS trials were associated with "abdominal pains, brainstem-related cardio-respiratory arrest " (active treatment, dose unknown) and "unknown;" (IFN 3 MIU TIW). The non-suicide SPMS deaths were associated with intracerebral hemorrhage and subarachnoid hemorrhage (blinded treatment assignment).

There were 15 dropouts due to adverse events from trials in RRMS other than --------. Among the 4 dropouts due to serious adverse events (the ones for which there is an event description) 2 were due to depression (one with a suicide attempt), 1 was due to dysphasia, and 1 was due to an increase in dizziness. There were 37 dropouts due to adverse events in secondary progressive MS. Among the 17 dropouts due to serious adverse events, for which there is an event description, at least 5 were due to depression-related events, including 1 suicide attempt. Other causes were diverse, with no single predominating cause. There were 4 dropouts due to adverse events in other MS forms and demyelinating disease; none were due to serious adverse experiences and there are no additional data on them.

Including deaths (and including serious adverse experiences leading to dropout discussed immediately above, but excluding scheduled hospitalizations for surgery planned before the onset of a trial), there were 23 serious adverse experiences in RRMS trials other than --------, 151 in secondary progressive MS, and 4 in other demyelinating diseases. As Table 32 shows, in RRMS the treatment attribution is unknown except in 3 cases with IFN treatment; cases were of diverse natures.

Table 32. Serious adverse events in RRMS trials other than -------- (excludes hospitalizations for planned surgery; "x" is IFN at unknown dose; "-" is blinded)

Body System	Description (number of cases)	Treatment
Body as a whole-general	Sudden death of unknown cause	IFN 3 MIU
CNS and PNS	Dysphasia , cognitive impairment Increased dizziness Deterioration of condition Brain stem lesions leading to death	IFN 3 MIU IFN 1.5 MIU x x
GI system	Nausea and vomiting Gastrointestinal flu Gastroenteritis Acute appendicitis	- - - -
Liver and biliary	Cholelithiasis Chronic cholecystitis	- -
Neoplasms	ovarian cystadenocarcinoma	-
Psychiatric	severe depression (2)	-
Red blood cell	anemia	x
Respiratory system	Pneumonia Pneumonia; empyema	IFN 9 MIU x
Urinary system	Suprapubic catheterization Acute urinary retention Acute pyelonephritis	- x -
Vascular (extracardiac)	popliteal deep venous thrombosis	-
White cell and RES	severe lymphopenia	-
Surgery	surgery for trigeminal neuralgia	-

There were 151 serious adverse event reports in SPMS trials, excluding planned hospitalizations (and, in trial -------, hospitalizations for MS exacerbations). The treatment assignment for all these events is still blinded. The spectrum of disorders is similar to that for ---------. The body systems most affected are: psychiatric (21 cases:14 severe depressions, 2 completed suicides, 3 suicide attempts, 1 drug overdose, and 1 anorexia); urinary (20 events, among which were 11 infections and 4 cases of urinary retention); and trauma (19 events, among which were 10 fractures). The overall rate of serious adverse event rates for trials in SPMS was twice that for RRMS: 11.1 for each 100 subject-years of exposure in SPMS, compared to 5.4 in trials in RRMS (these numbers include placebo subjects). Other trends, such as the differences in proportion of organ systems affected, will not be discussed here.

There have been 4 serious adverse event reports in trials of the MS-related disorders "clinically probable or laboratory supported definite MS" and chronic inflammatory demyelinating polyneuropathy. The treatment assignment for these events is still blinded. Two of these were psychiatric (emotional lability and anxiety); 1 was a case of elevated transaminases, and 1 was a case of purulent tonsillitis.

Summary

The lack of knowledge of treatment assignment in most of the reported serious adverse events renders attribution to IFN problematic. The spectrum of disorders of concern was similar to that seen in the controlled trial --------. It should be noted that Serono reports that the incidence of appendicitis was higher than that seen in the general population (8 cases as opposed to the expected 3). The explanation for this is beyond the scope of this review. In general, serious adverse events in secondary progressive forms of MS were of diverse natures, with an increase in some types (e.g., in psychiatric and urinary systems).

Trial --------- represents 63% of all subject-years (1061/1693, including placebo subjects) in trials of RRMS conducted by Serono, and as a completed controlled trial represents the best assessment of the safety (and efficacy) of the product to date. The integrated summary of safety does not contain reasons for extra concern.

120-DAY SAFETY UPDATE

The 120-day safety update for Rebif comprises data collected from January 1, 1997 to February 28, 1998. Three new trials are included in the data base. They are all open-label trials in RRMS, involving 729 new subjects, mostly at 22 mcg once a week SC, and representing 243 additional subject-years of exposure in those trials alone. The reporting period accounted for 15 clinical trials in MS (including ----------) and related indications. Including control subjects, this represents 1254 subject-years of observation in RRMS, 978 subject-years in secondary progressive MS, and 289 subject-years in early-onset MS. Taking all trials during the reporting period, there were 3 suicides, 2 on active treatment and one in a blinded trial; the other deaths were attributed to lymphoma in the brain and to probable pulmonary hemorrhage following convulsions. There were 71 serious adverse events during the reporting period in RRMS, 96 in secondary progressive MS, and 9 in other MS disease forms and demyelinating diseases. The spectrum of disorders for RRMS patients is similar to that in ---------; trauma was the largest category of disorder (14 serious adverse events), followed by psychiatric events (9, including an estimated 7 serious depressions). The distribution of organ systems affected in SPMS trials was very similar to that reported in the integrated summary of safety.

Incidental to this review but of possible importance to the understanding of IFN treatment in other MS indications was the greater incidence of serious adverse events, including serious depression, in the secondary progressive indication than in the RRMS (17 serious depressions in SPMS as opposed to 7 in RRMS; numbers estimated from Serono's tabulation of adverse event descriptions). Represented as rates per 100 subject-years of observation, the two highest categories in secondary progressive MS were "psychiatric disorders" and "trauma" (1.6 and 1.3, respectively), each one representing a higher rate than the highest rates in RRMS (0.7, for psychiatric disorders and surgery each).

Summary

The 120-day safety update does not change the overall impression of the safety profile of IFN-β1a as used in RRMS gleaned from the summary of safety in the controlled trial ----------.

RECOMMENDATIONS

1) Both dosages of Rebif are approvable for relapsing-remitting patients, since in this category of patients both caused clinically significant changes in most of the primary and secondary outcome measures used in the controlled trial, and neither was associated with excessive numbers of dose-limiting toxicities. However, since efficacy did not differ substantially between the two doses, and the incidences of troubling adverse events (for example, cytopenias and hepatic enzyme abnormalities) was notably larger for the high dose, patients should be advised to initiate dosing at 22 mcg TIW.

2) Since the safety and efficacy of this product were determined in subjects who initiated treatment with a dose escalation, instructions for use should include similar instructions to those given the subjects in the trial, and a statement should be included to the effect that safety and efficacy have only been studied under this dose regimen.
3) The primary outcome as analyzed by Serono is acceptable for inclusion in package labeling, as different manipulations of the method of its analysis did not change the measured extent of benefit, and the benefit was clinically meaningful.
4) Secondary outcomes that might be included in the package labeling are:
   - Time to first exacerbation
   - Duration of exacerbations
   - Time to 3-month progression in disability
   - Steroid use rate for MS symptoms
   - Percent change in PD/T2 lesion area and in lesion activity
5) Secondary outcomes not appropriate for labeling are the following, for indicated reasons:
   - Time to second exacerbation: This has marginal added utility.
   - Severity of exacerbations: There is no evidence that the effect of IFN on the severity of exacerbations is independent of its effect on the incidence of exacerbations of all severities.
   - Hospitalizations for MS: These differences were of marginal significance for the low dose, although trending in the same direction toward benefit.
6) Labeling should include the precautions and warnings applicable for the other marketed IFNs (Avonex and Betaseron).
7) Approval of Rebif for categories of MS patients other than those with the relapsing-remitting variety should be contingent upon direct testing of such patients in well-controlled, randomized, double-blinded clinical trials, with prospectively defined, widely accepted entrance criteria.
8) MRI scans should be subjected to further review to eliminate concerns over inter-rater variability.

REFERENCES

1) Poser, C.M., Paty, D.W., Scheinberg, L., McDonald, I., Davis, F.A., Ebers, G.C., Johnson, K.P., Sibley, W.A., Silberberg, D.H., and Tourtellotte, W.W. (1983) New diagnostic criteria for multiple sclerosis: Guidelines for research protocols. Ann. Neurol. 13: 227-231.
2) Barnes, D., Hughes, R.A.C., Morris, R.W., Wade-Jones, O., Brown, P., Britton, T., Francis, D.A., Perkin, G.D., Rudge, P., Swash, M., Katifi, H., Farmer, S., and Frankel, J. (1997) Randomised trial of oral and intravenous methylprednisolone in acute relapses of multiple sclerosis. Lancet 349: 902-906.

APPENDIX

Blinding questionnaire:Subjects' and investigators'opinions concerning which treatment they received

At the end of the trial subjects were asked, “In your opinion which treatment did you receive during the previous 24 months of the study?” Five hundred and seventeen subject questionnaires were filled in. The responses indicate that most subjects on IFN suspected that they had received active treatment.

Table 33. Subjects’ responses to blinding questionnaire concerning which treatment they received

Subject opinion	Actual treatment		
	placebo	IFN 22 mcg TIW	IFN 44 mcg TIW
placebo	59	22	9
active treatment	65	102	132
I do not know	50	46	32
total	174	170	173

Evaluating physicians were asked the question, “In your opinion which treatment did this patient receive during the previous 24 months of the study?” Five hundred and twenty-one evaluating physician questionnaires were filled in.

Table 34. Evaluating physicians’ responses to blinding questionnaire concerning which treatment their patients received

Evaluating physician opinion	Actual treatment		
	placebo	IFN 22 mcg TIW	IFN 44 mcg TIW
placebo	43	30	19
active treatment	46	48	66
I do not know	86	94	89
total	175	172	174